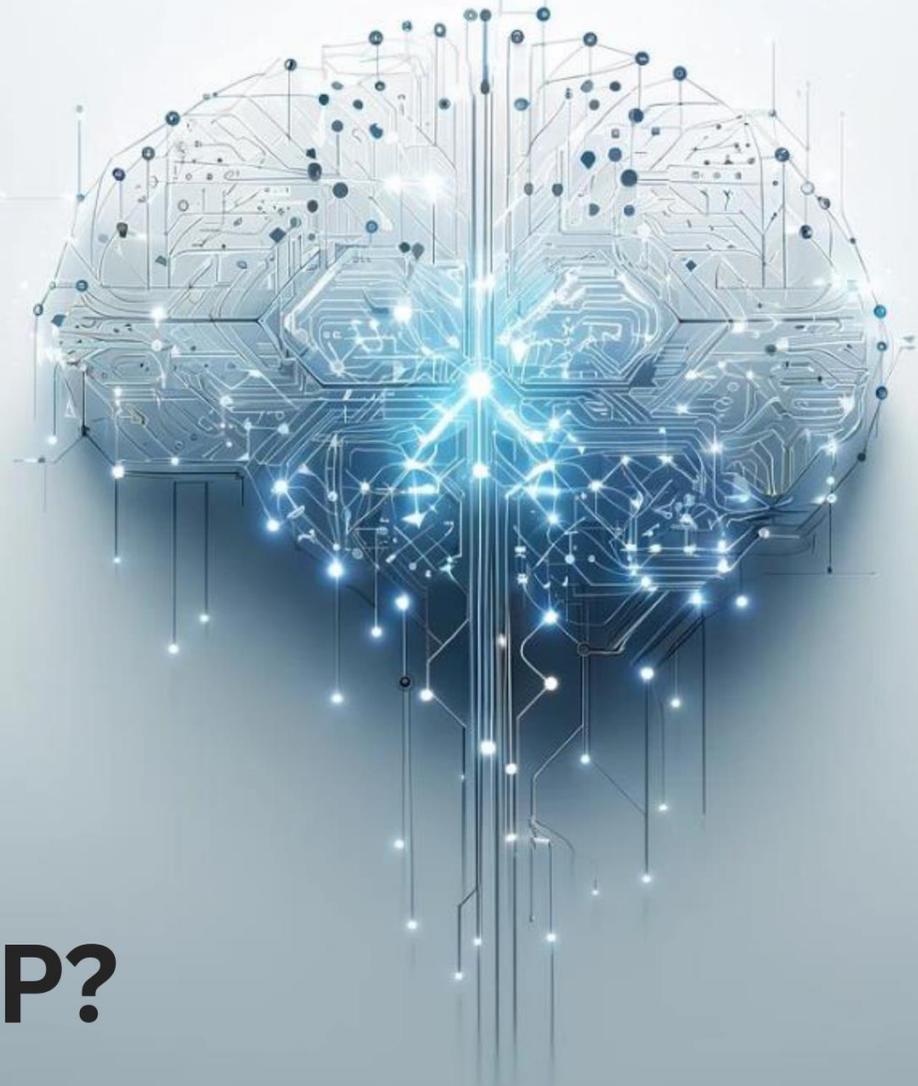


 **APRAM**

ECTA

Caring for a balance of IP rights



Co-organised with:


PANTHÉON SORBONNE
UNIVERSITÉ PARIS 1
ÉCOLE DE DROIT
DE LA SORBONNE

AI & IP: A LOVE-HATE RELATIONSHIP?

Sponsored by:

fovea^{IP}
The sharpest Eye on IP

 **E BRAND**

**APRAM-ECTA[®] ETREAT JOINT EVENT in Paris, Panthéon Sorbonne
3 April 2025, 9:00-17:00 followed by a networking cocktail**

AI & IP: A LOVE-HATE RELATIONSHIP?

PROGRAM

3 APRIL 2025, PANTHEON SORBONNE AMPHI TURGOT, 17 RUE DE LA SORBONNE, PARIS

Co-organised with



Sponsored by



C



09:00 ARRIVAL OF PARTICIPANTS

09:15 WELCOME SPEECHES & SPONSORS PRESENTATION

09:30 Panel 1: A short overview of technical aspects of AI and different perspectives on protection

Moderator: EDOUARD FORTUNET, Partner, Jones Day, FR
JEAN-MARC DELTORN, Professor, CEIPI, FR

10:10 Panel 2: AI as a source of both creation and infringement. How has it been tackled: case law from different jurisdictions

PIERRE MASSOT, Partner, Founder of ARENAIRE Lawfirm, FR
CHARLES-ANTOINE JOLY, Partner, DS Avocats, FR
ALAIN STROWEL, Professor, UCLouvain, Partner Pierstone Brussels, BE

11:30 Q&A

12:00 – 13:30 BREAK (Please note that no lunch is organised for participants)

13:45 OPENING OF THE AFTERNOON SESSION

13:55 Panel 3: Data scraping to feed AI models, data protection and other IP issues

Moderator: DELIA BELCIU, Partner, DB Law Office, RO

Rights reservations pursuant to Art. 4(3) DSM Directive and the policy of GPAI model providers to identify and comply with them

Introduction by EDOUARD TREPPOZ, Professor, Panthéon - Sorbonne Paris I University, FR
Presentation by ALEXANDER PEUKERT, Professor of Civil Law and Commercial Law, Goethe University Frankfurt, DE

Discussion: IP aspects linked to Article 53 of the AI Act transparency obligation and how to prove that AI providers have used your work

ALEXANDRA BENSAMOUN, Professor, Université Paris Saclay, FR
THOMAS BOUVET, Partner, Jones Day, FR

15:30 Panel 4: Implications of AI and industry perspectives

Moderator: NICOLAS MOREAU, Partner, Bignon Lebray, FR
DAVID EL SAYEGH, Deputy Director-General, SACEM, FR
ANGELIQUE DE SIMENCOURT, Deputy Director IP, CHRISTIAN DIOR COUTURE, FR
BENJAMIN BELLAMY, Business Development Manager for AI Solutions, LINAGORA, FR

16:45 Q&A

17:00 – 19:00 CLOSING COCKTAIL AT PANTHEON, 133 RUE SAINT-JACQUES, PARIS

THANK YOU TO



&

OUR SPONSORS



&





EBRAND

www.ebrand.com

Ask the bear
06 80 48 68 06
pberecz@ebrand.com

Ours né sous IA



**Détections, analyses et suppressions
des atteintes aux marques
sur les *places de marché*,
les *sites Internet*
et *e-commerce*,
les *réseaux sociaux*,
les *app stores*,
le *darkweb*,
et dans les *noms
de domaine...***

Notre
mission :
Nettoyer
l'Internet



*Demandez
à un ours
de nettoyer
l'Internet !*

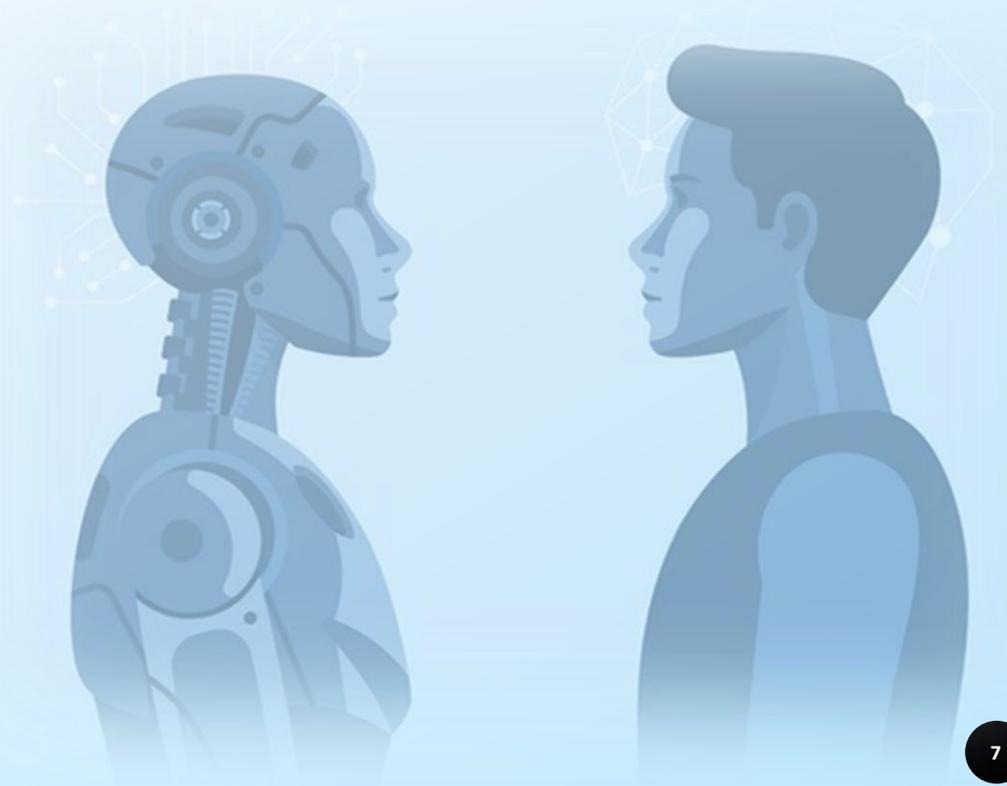


Demandez
à un ours
de nettoyer
l'Internet!

AI & IP: A LOVE-HATE RELATIONSHIP?

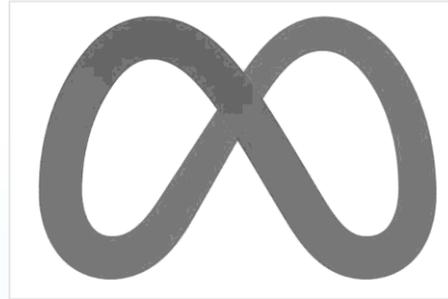
Artificial Intelligence is by Fovea IP analysts and clients where it brings

- more **reliability**
- more **efficiency** and, therefore
- **reduced costs** and **faster delivery times**



Logo searching

EU27 | classes 9, 35, 42



Human Brain

+ Vienna classes 24.17.08, 27.03.15,
27.05.21, 27.99.13



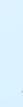
135 762 active trademarks



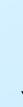
5 days

Artificial Intelligence

No Vienna classes



2 120 979 active trademarks



3 seconds

AI Figurative Similarity

REGISTERS

Trade Marks (28/196): Austria, Benelux, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, EUIPO, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Malta, ...

Please upload your image, available formats (JPG, JPEG, PNG), size (0.008MB - 2MB), pixels (100px, 1536px)



BROWSE

Nice Classes 

Claimed Countries (WIPO)

Select one or more values

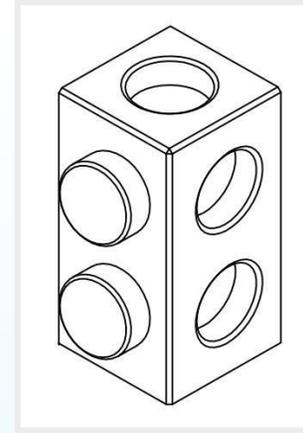
Reference

 SEARCH

 CLEAR

Industrial design searching

11 countries | subclasses 0901 & 0903



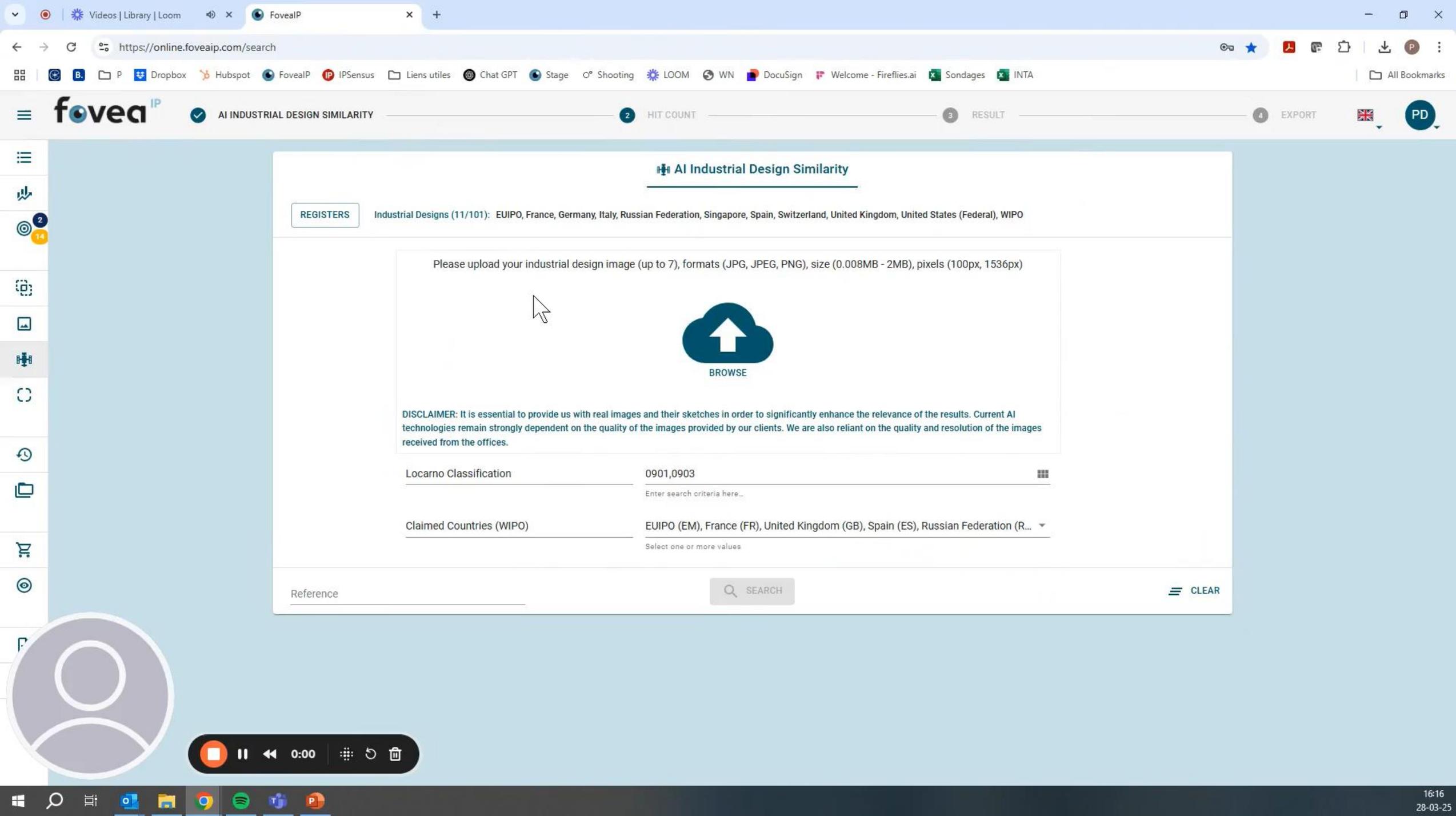
Human Brain

Artificial Intelligence

319 423 active & inactive industrial designs

8 days

6 seconds



AI Industrial Design Similarity

REGISTERS

Industrial Designs (11/101): EUIPO, France, Germany, Italy, Russian Federation, Singapore, Spain, Switzerland, United Kingdom, United States (Federal), WIPO

Please upload your industrial design image (up to 7), formats (JPG, JPEG, PNG), size (0.008MB - 2MB), pixels (100px, 1536px)



DISCLAIMER: It is essential to provide us with real images and their sketches in order to significantly enhance the relevance of the results. Current AI technologies remain strongly dependent on the quality of the images provided by our clients. We are also reliant on the quality and resolution of the images received from the offices.

Locarno Classification

Enter search criteria here...

Claimed Countries (WIPO)

Select one or more values

Reference

SEARCH

CLEAR



Fovea IP is also **200 human brains** (and growing) to

- **Review** A.I. work (logos, industrial designs)
- **Perform**
 - Word mark searches (with the help of algorithms)
 - Watching (with the help of algorithms)
 - In-use searches
 - Brand awareness surveys
- **Manage** +200 million trademarks, +23 million industrial designs, +170 million patents, +600 000 000 domain names, and a lot more...

Contact us for more info:

Pauline Déom & Marc Tena

Commercial Direction French-speaking countries

pauline.deom@foveaip.com

+32 471 48 18 55

marc.tena@foveaip.com

+33 6 12 11 40 39



APRAM

ECTA

Caring for a balance of IP rights

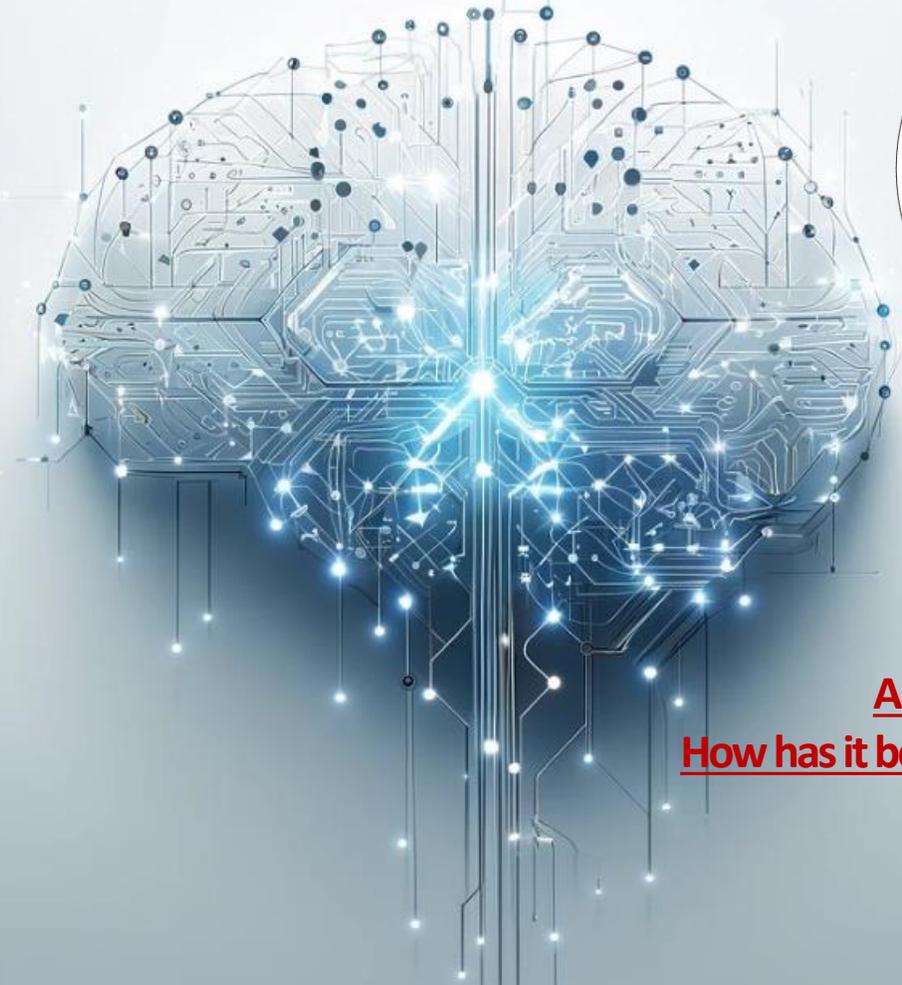
Co-organised with:



Sponsored by:



AI & IP:
A LOVE-HATE RELATIONSHIP?



Moderator:
EDOUARD FORTUNET
Jones Day, FR

A short overview of technical aspects of AI and different perspectives on Protection

Prof. JEAN-MARC DELTORN
CEIPI, FR

AI as a source of both creation and infringement. How has it been tackled: a se law from different jurisdictions

PIERRE MASSOT
ARENAIRE Lawfirm, FR
CHARLES-ANTOINE JOLY
DS Avocats, FR
Prof. ALAIN STROWEL
Saint-Louis (Brussels); UCLouvain, BE

APRAM-ECTA[®] ETREAT JOINT EVENT (Paris, 3 April 2025, 9:00-17:00)

A very brief introduction to Artificial Intelligence

Jean-Marc Deltorn

CEIPI

University of Strasbourg

DENoTE, MANTRA

What is AI?

Can we attempt to define “artificial intelligence”?



Dartmouth: "IA"
(birth of the term)

"Deep belief network pretraining (Geoff Hinton): the renaissance of deep neural networks (DNNs)

First industrial developments: Waymo, Uber, Tesla (etc...) (semi)-autonomous vehicles

Deepmind (Alpha Go, Alpha Zero), etc...

Acceleration: Dall-E, StableDiffusion, Suno, Foundationmodels, LLMs, GPT-o3, DeepSeek
Nobel prizes,
The emergence of regulation...
the ubiquity of AI

1950

1970-2000

2011-2012

2017

2020-2021



Alan Turing
"Can machine think?"

1955

2006

2016

2018-2019

2022-2025

symbolic and sub-symbolic approaches (summers and winters of AI)

DNN at scale: Andrew Ng and Jeff Dean Google Brain "unsupervised recognition

The renaissance of generative AI: "generative adversarial networks" ("GANs")... and Deepfakes

openAI GPT2: "too dangerous to share", AlphaFold: AI and innovation

- **John McCarthy's Definition (1955):** "the science and engineering of making intelligent machines."
- **Stuart Russell and Peter Norvig:** "the study of agents that perceive their environment and take actions that maximize their chances of achieving their goals." ("Artificial Intelligence: A Modern Approach")
- **Margaret A. Boden (philosopher):** "the use of computer programs and programming techniques to cast light on the principles of intelligence in general and human thought in particular."
- **EU AI Act Definition:** An AI system is a machine-based system designed to operate with varying levels of autonomy, capable of generating outputs such as predictions, recommendations, or decisions that influence physical or virtual environments.

What is AI?

An “umbrella term”, a elusive goal, ill-defined...

In practice: “AI” involves creating machines (computer programs) capable of performing tasks that typically require human intelligence, such as learning from examples, producing new artefacts and problem-solving.

Is AI new?

Not really : AI has progressed from early neural networks (a few artificial neurons) to very large models deep learning models with trillions of parameters (with many other techniques in between, incl. symbolic, hybrid, probabilistic, evolutionary, etc.), reflecting continuous advancements in algorithmic solutions, technology (CPU, GPU, TPU) and access to data

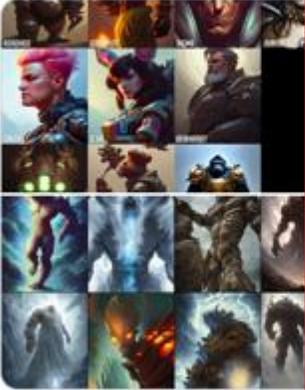
Why such a hype?

Is this a real technological / societal disruption or just speculation?

Is it here to stay?

RJ Palmer
@arvalis · Follow

A new AI image generator is making art that looks 100% like an artist I am extremely close to.



3:16 AM · Aug 14, 2022

[Read the full conversation on](#)

89K Reply Share

[Read 1.6](#)

FEB. 27 / MARCH 6, 2023

Hi ChatGPT

Hello! How can I assist you today?

What do you think of a TIME cover story with the title, "The AI Arms Race Is Changing Everything?"

As an AI language model, I don't have personal opinions or beliefs, but I can provide some context and information on the subject.

"The AI Arms Race is Changing Everything" could be an attention-grabbing and thought-provoking title for a cover story on AI. A cover story in a widely-read publication like TIME could help raise public awareness about the potential risks and benefits of this trend, and stimulate discussions on how to ensure that AI is developed and deployed in a responsible and ethical manner.

It's important to note that the title of the story could also be interpreted as sensationalist or alarmist, and may not accurately reflect the complexities

Sincarnate 08/26/2022

Hi everyone! After a one month hiatus, I have returned with an exciting announcement about my personal project I've made using Midjourney. I have created a special prompt that I will be publishing at a later date, I have generated a lot of images using it, and after many weeks of fine tuning and curation, I chose my top 3 and had them printed on canvas after being approved by Gigapixel A.I. I entered into the Colorado State Fair fine arts (digital arts category). I won first place. Here are the results. I've made a statement using Midjourney in a competitive manner and I cannot be more excited about having won with my favorite piece: "Opera Spatial"



by
ANDREW R. CHOW
& BILLY PERRIGO
(humans)

time.com

Finance worker pays out video call with deepfake officer'

By Heather Chen and Kathleen Magramo, CNN
2 minute read · Published 2:31 AM EST, Sun February 4, 2024



Nobel prizes 'They don't just fall out awards highlight Britain'

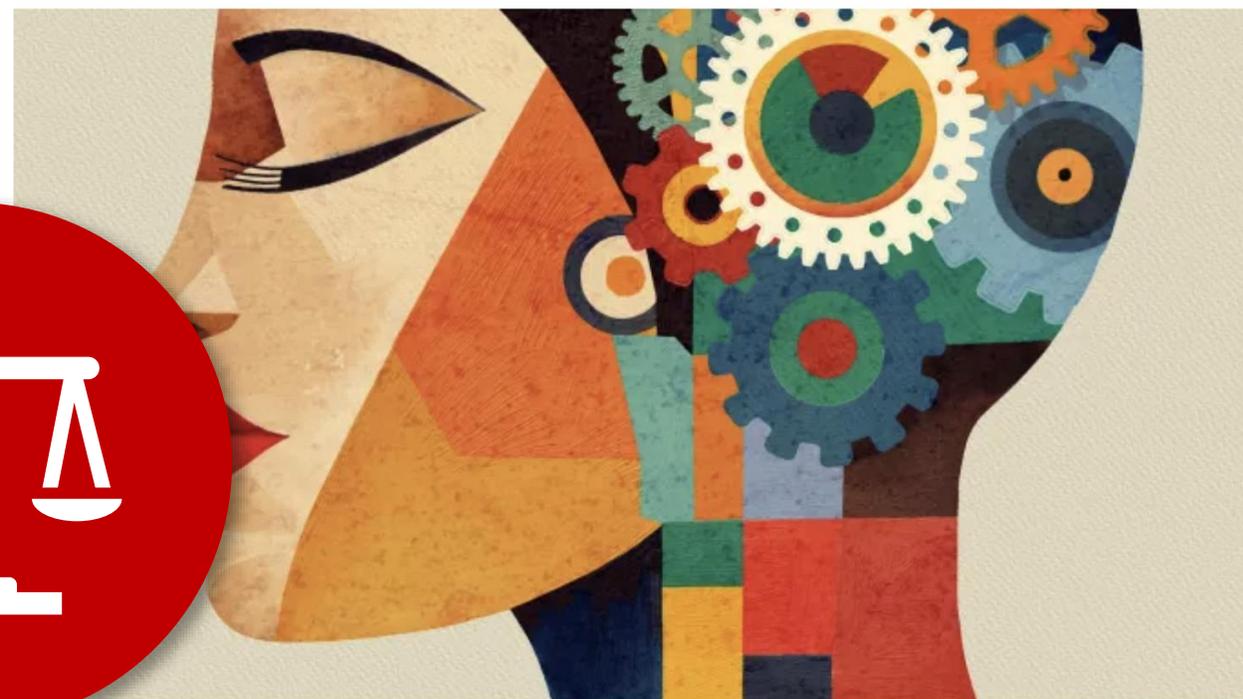
Recognition for Demis Hassabis and Google moment when important ingredients c



Artists Score Major Win in Copyright Case Against AI Art Generators

The court declined to dismiss copyright infringement claims against the AI companies. The order could implicate other firms that used Stable Diffusion, the AI model at issue in the case.

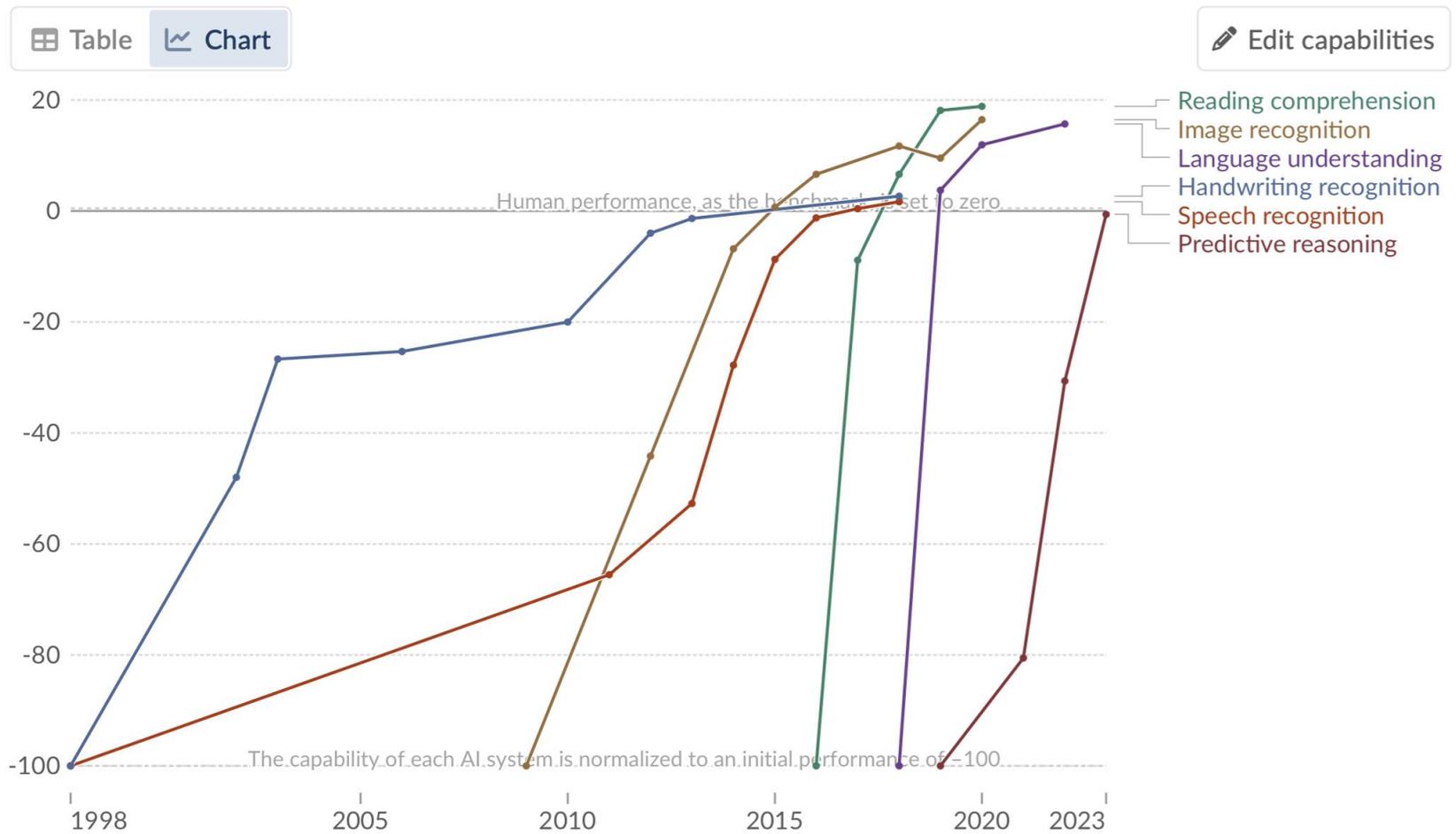
BY WINSTON CHO AUGUST 13, 2024 1:09PM



BORIS SV / GETTY IMAGES

Test scores of AI systems on various capabilities relative to human performance

Within each domain, the initial performance of the AI is set to -100. Human performance is used as a baseline, set to zero. When the AI's performance crosses the zero line, it scored more points than humans.



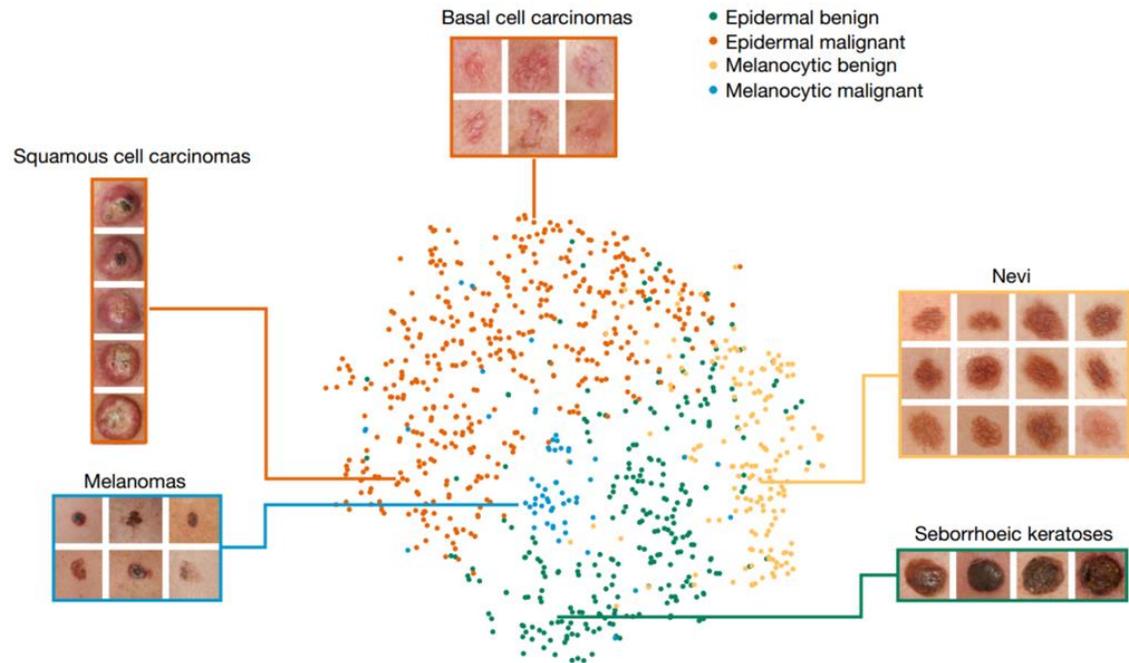
Data source: Kiela et al. (2023) – [Learn more about this data](#)

OurWorldinData.org/artificial-intelligence | CC BY

Note: For each capability, the first year always shows a baseline of -100, even if better performance was recorded later that year.



Not just academic... but real-world applications



Esteva et al. 2017 "Dermatologist-level classification of skin cancer with deep neural networks" Nature 542, 115-118

naturemedicine

Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

[nature](#) > [nature medicine](#) > [articles](#) > article

Article | Published: 17 May 2024

Prediction of DNA methylation-based tumor types from histopathology in central nervous system tumors with deep learning

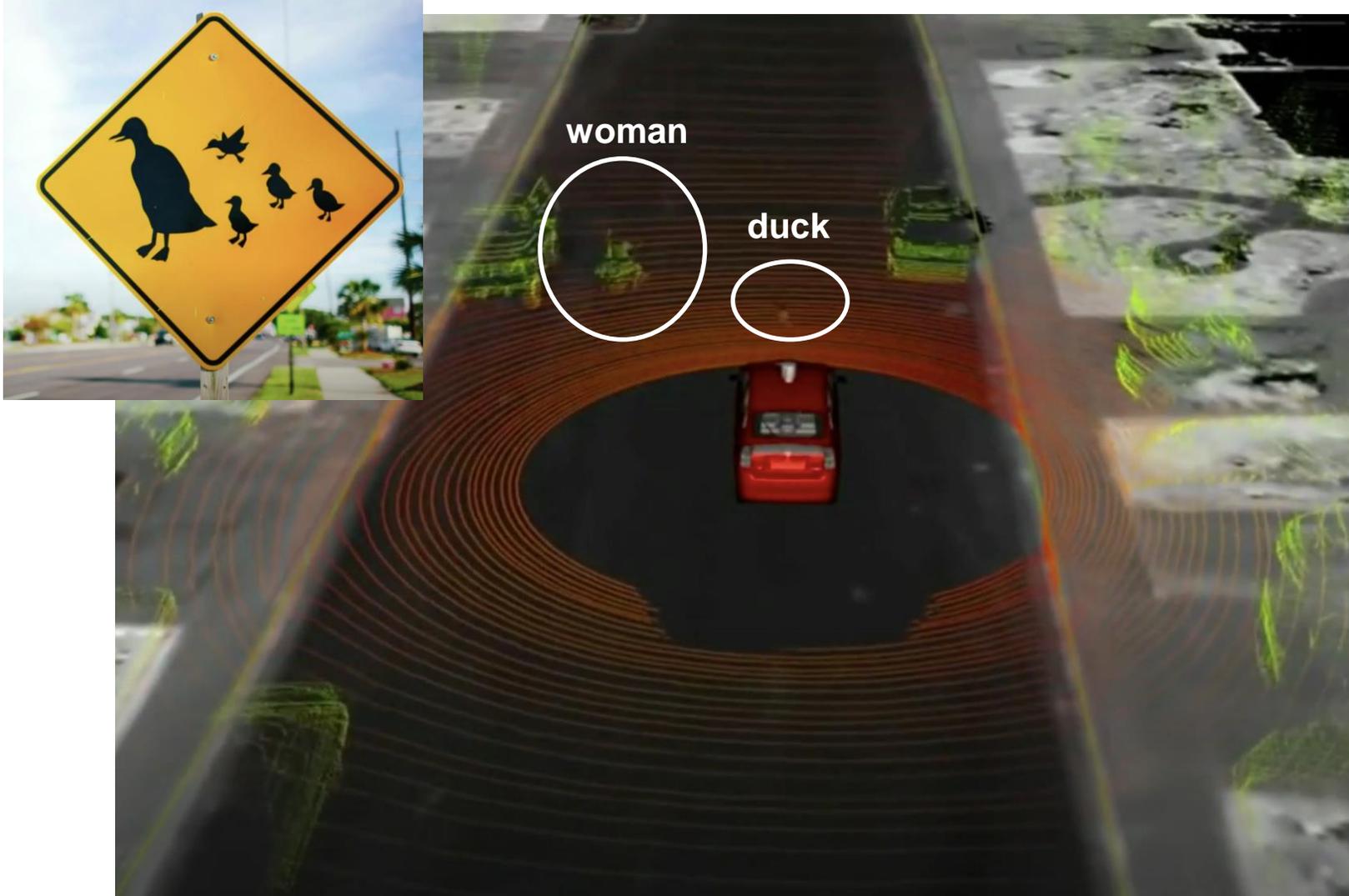
[Danh-Tai Hoang](#), [Eldad D. Shulman](#), [Rust Turakulov](#), [Zied Abdullaev](#), [Omkar Singh](#), [Emma M. Campagnolo](#), [H. Lalchungnunga](#), [Eric A. Stone](#), [MacLean P. Nasrallah](#), [Eytan Ruppin](#) ✉ & [Kenneth Aldape](#) ✉

Nature Medicine 30, 1952–1961 (2024) | [Cite this article](#)

7385 Accesses | 5 Citations | 649 Altmetric | [Metrics](#)

Abstract a ten-class model trained on an internal dataset of 1,796 patients, we predict the tumor categories in three independent external test datasets including 2,156 patients, achieving an overall accuracy of 95%

Reality checks: coping with edge cases & safety



“Just a couple of months ago, our vehicles were driving through Mountain View, and this is what we encountered : this is a woman in an electric wheelchair, chasing a duck in circles on the road...”

(Chris Urmson, formerly with Waymo)

AI Incidents

The AI Incident Database (AIID) tracks instances of ethical misuse of AI, such as autonomous cars causing pedestrian fatalities or facial recognition systems leading to wrongful arrests.² As depicted in Figure 3.1.2, the number of AI incidents continues to climb annually. In 2023, 123 incidents were reported, a 32.3% increase from 2022. Since 2013, AI incidents have grown by over twentyfold.

The continuous increase in reported incidents likely arises from both greater integration of AI into real-world applications and heightened awareness of its potential for ethical misuse. However, it is important to note that as awareness grows, incident tracking and reporting also improve, indicating that earlier incidents may have been underreported.

Number of reported AI incidents, 2012–23

Source: AI Incident Database (AIID), 2023 | Chart: 2024 AI Index report

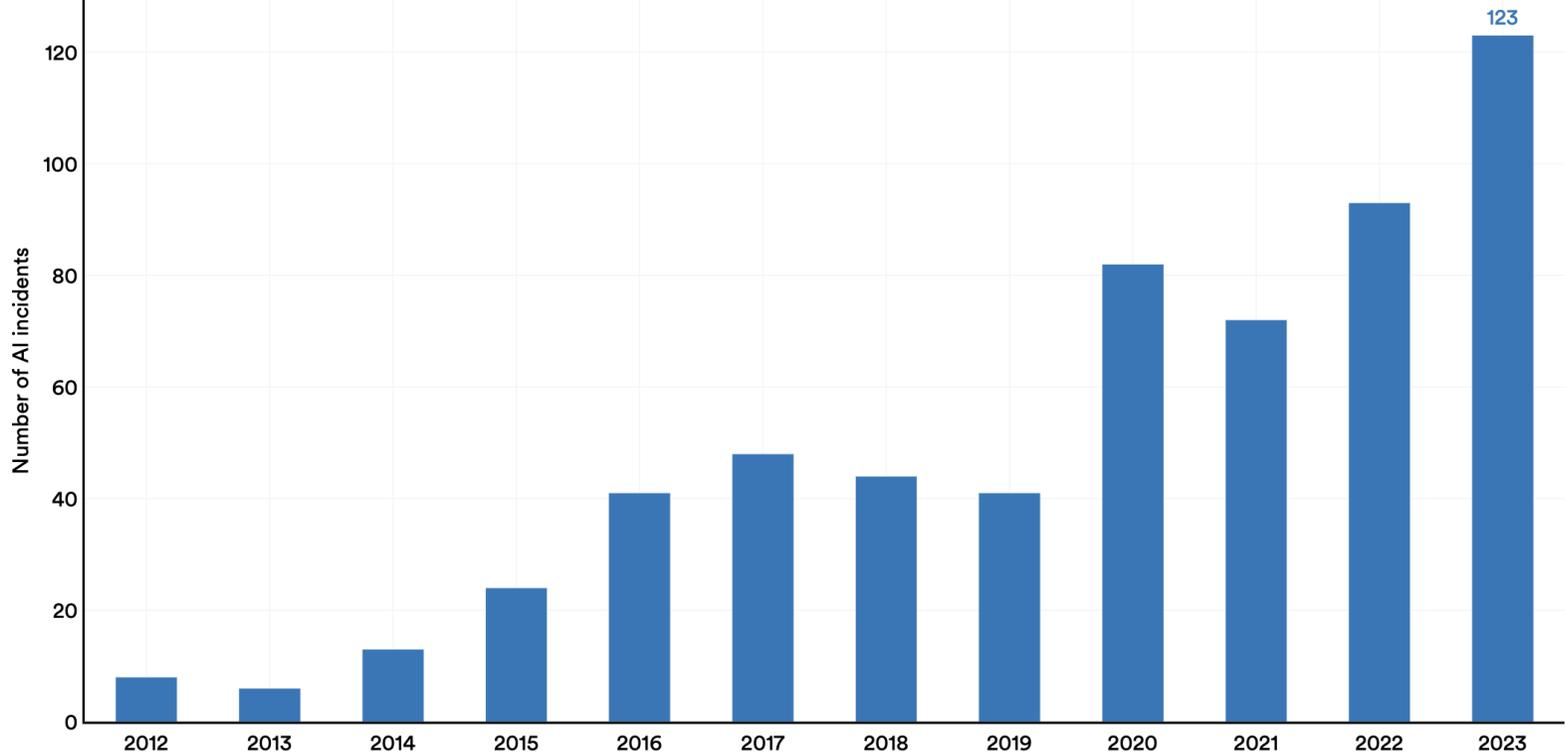
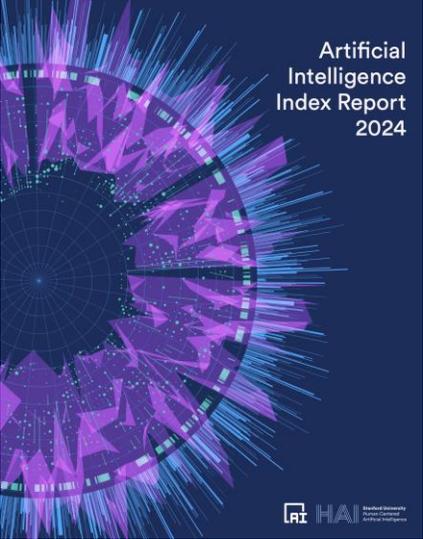


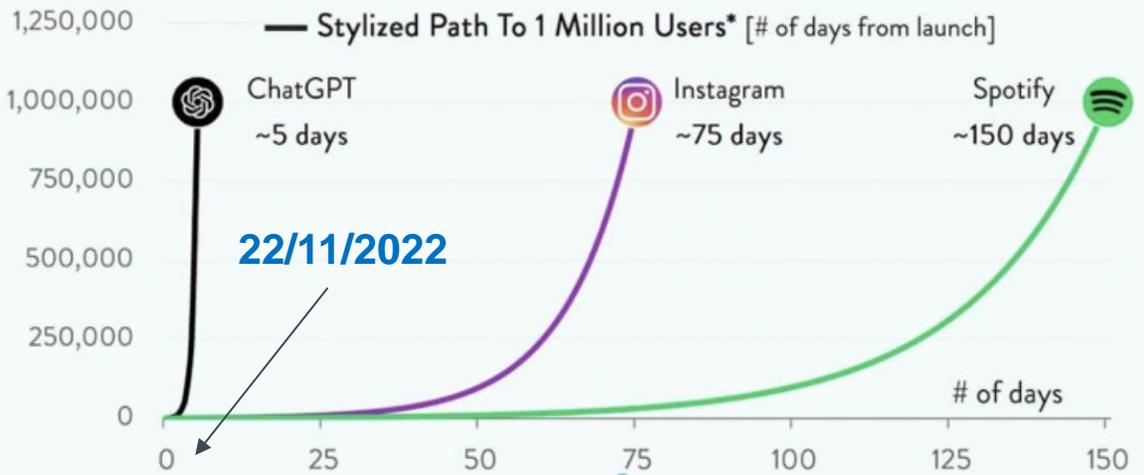
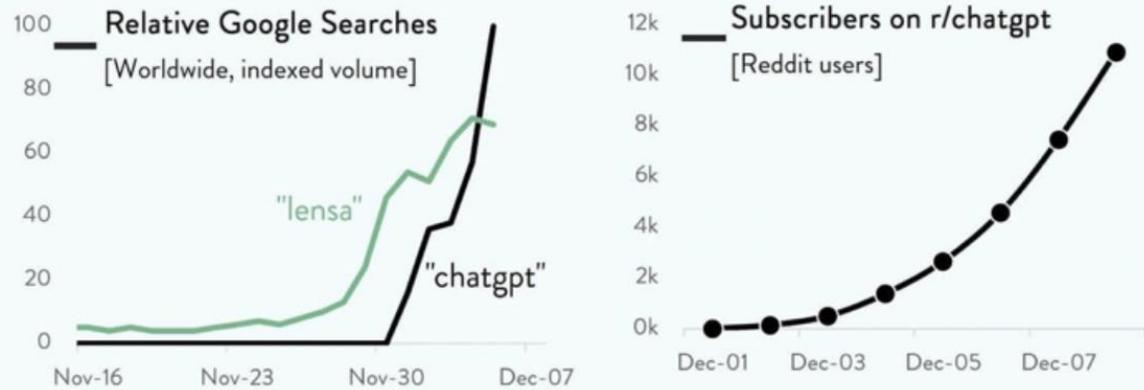
Figure 3.1.2



HAI - AI Index
Stanford 2024

Generative AI & LLMs

ChatGPT From OpenAI Is A Bot Taking The Tech World By Storm

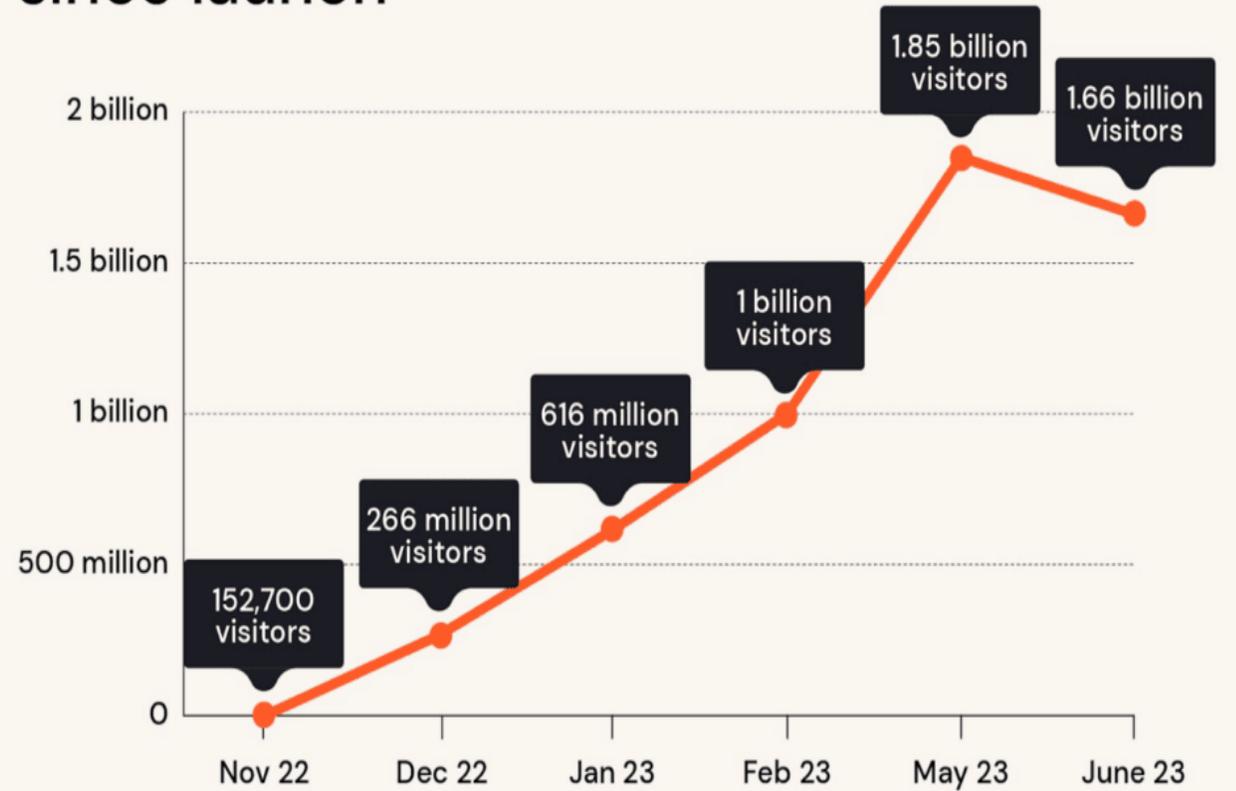


Sources: Google, Subredditstats, Media Reports

chartix *Path is stylized to the 1m milestone

CHATGPT STATISTICS

Change in ChatGPT website visitors since launch



Read the full report at tooltester.com/en/blog/chatgpt-statistics

tooltester

Why such a hype?

AI's rapid adoption across diverse industries is due to its ability to enhance efficiency, accuracy, and innovation (scientific discovery) in real-world applications.

AI is now delivering tangible benefits in a large spectrum of practical applications.

But, how does it work?

How can “machines learn”?

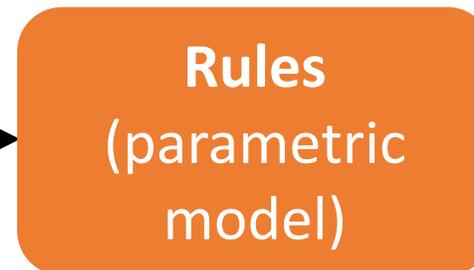
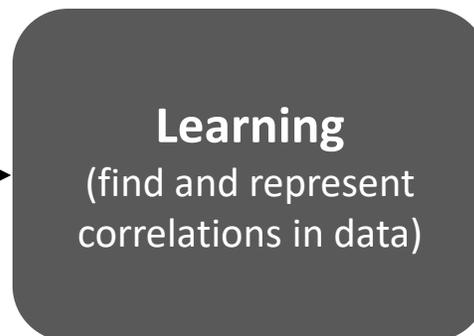
Artificial intelligence (AI): These are "machines whose behaviour would be considered intelligent if exhibited by a human". (also John McCarthy, 1955)

in practice :

- All theories, techniques and applications implemented to simulate cognitive processes
- (perception, **decision/production**, "understanding")

Expert systems

Humans must identify and set out rules in symbolic form, that are machine-interpretable (not easy!)



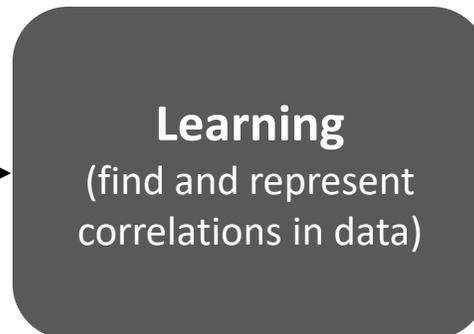
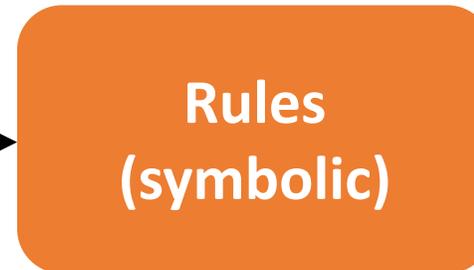
Artificial intelligence (AI): These are "machines whose behaviour would be considered intelligent if exhibited by a human". (also John McCarthy, 1955)

in practice :

- All theories, techniques and applications implemented to simulate cognitive processes
- (perception, **decision/production**, "understanding")

Expert systems

Humans must identify and set out rules in symbolic form, that are machine-interpretable (not easy!)



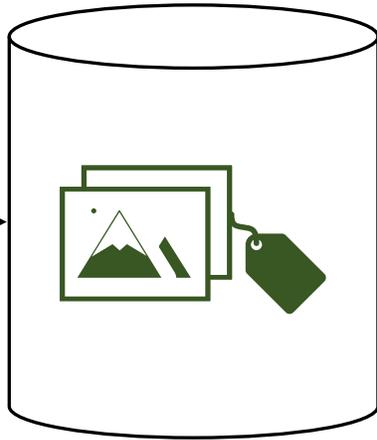
AI: from input to prediction

Start: definition of a problem (e.g. recognise classes of specific objects in images)

input data



annotators
(assign label to data)



training data

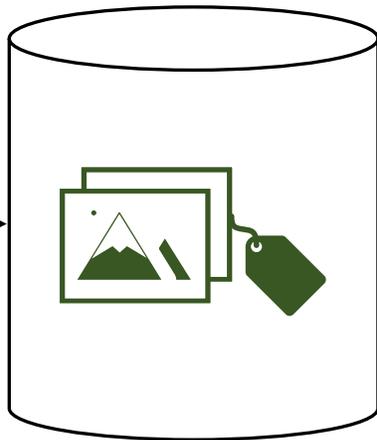
AI: from input to prediction

Start: definition of a problem (e.g. recognise classes of specific objects in images)

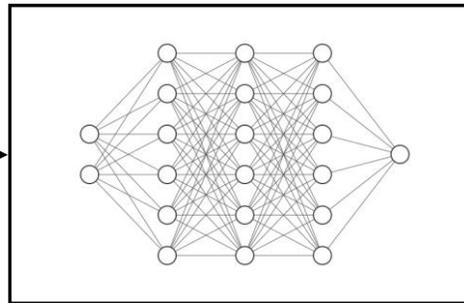
input data



annotators
(assign label to data)



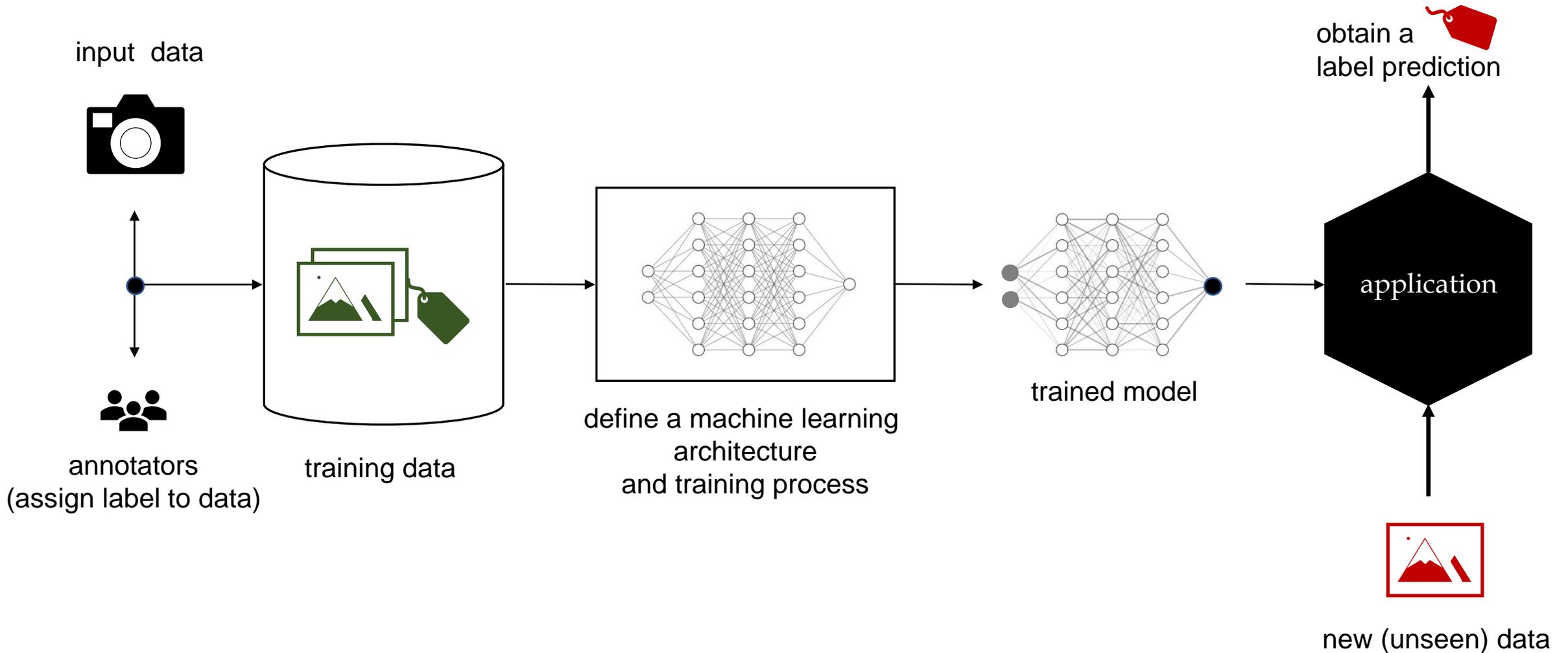
training data



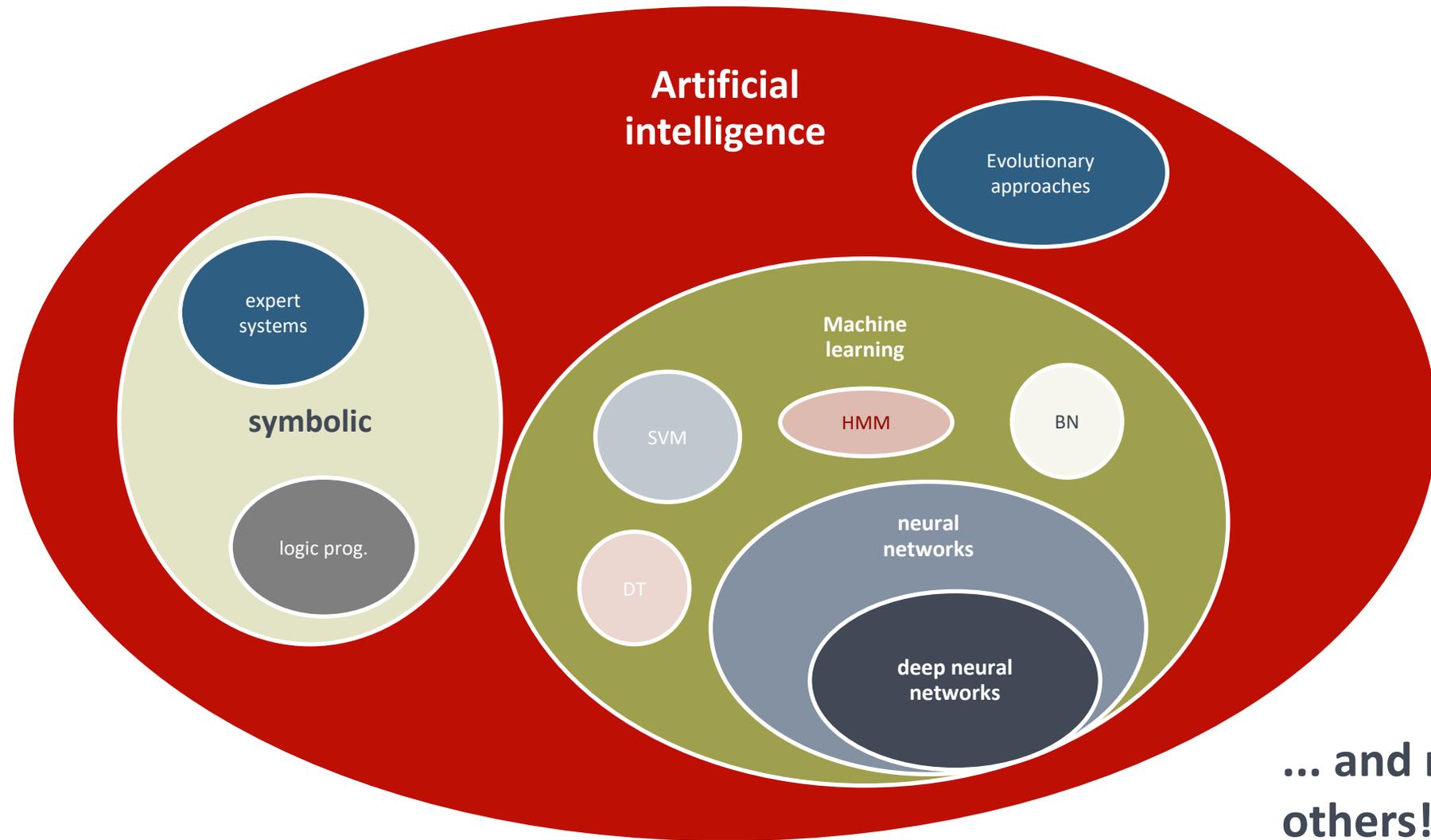
define a machine learning
architecture
and training process

AI: from input to prediction

Start: definition of a problem (e.g. recognise classes of specific objects in images)



AI: A (vast) array of techniques



... and many others!

But, how does it work?

AI systems learn from data through various methods, enabling them to make predictions, uncover patterns, or make decisions based on experience

How do you train an artificial neural network?

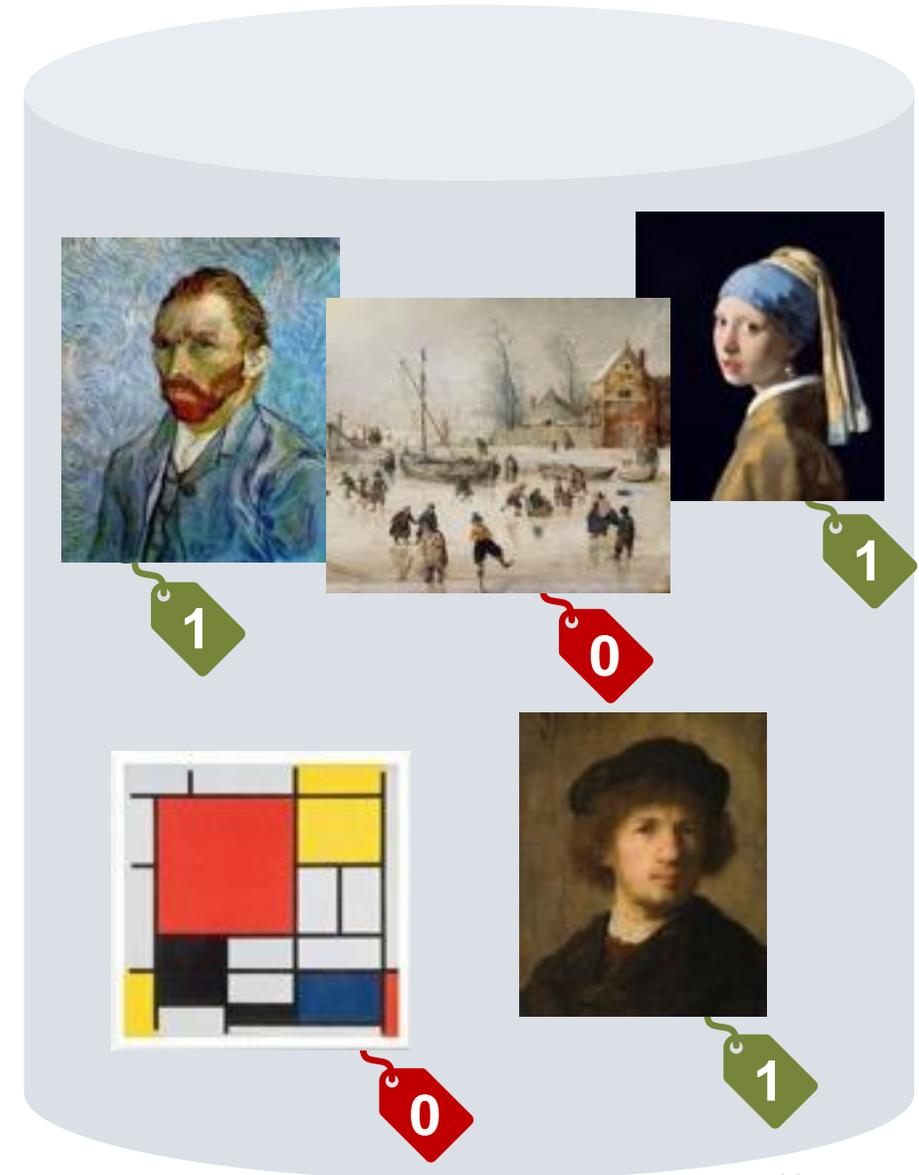
The problem: train a **classifier** to identify portraits among other types of work

How do we do it?

→ Using a **supervised learning** technique

(1) Choosing a data set

- we start with a set of labelled **training data** (in this case portraits and other types of painting).
- each image has been **labelled (by hand)** :
 - a "1" label for a portrait
 - a "0" label otherwise



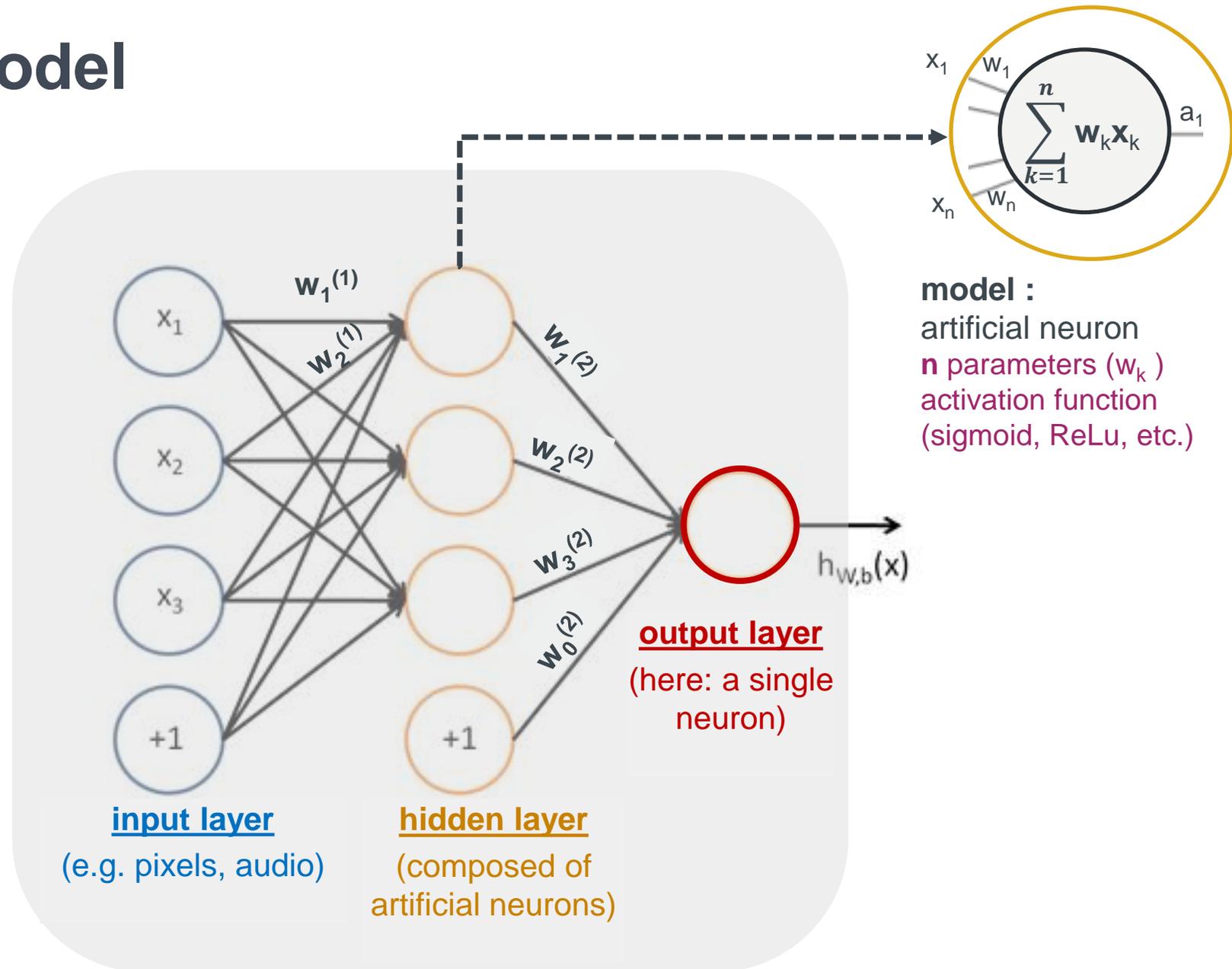
(2) Choosing a 'model

1. A neural network

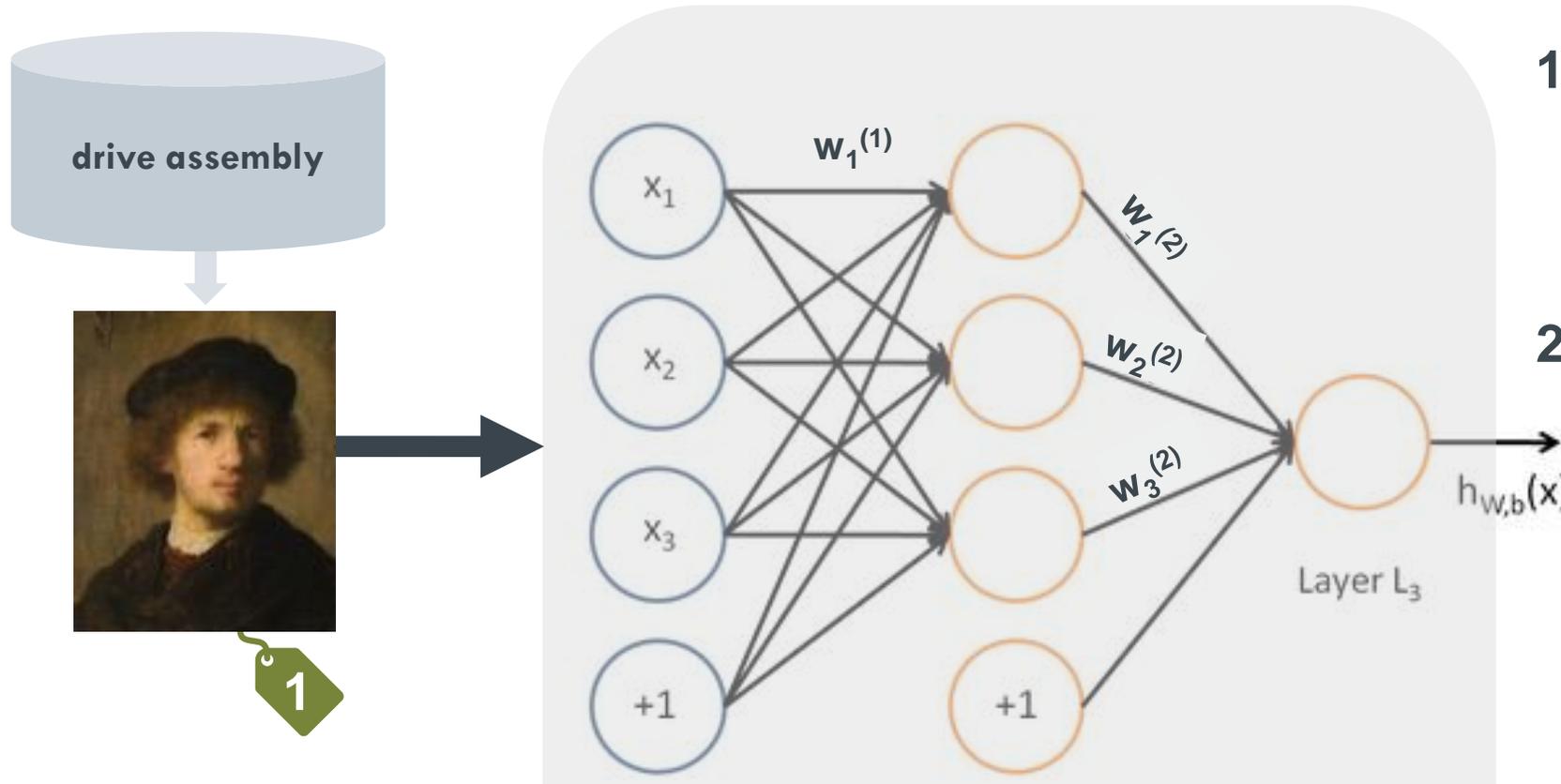
architecture is defined (number of layers, number of neurons per layer, types of activation functions, etc.).

Rq: The architecture may need to be adapted to the **type of input data** (text, sound, images, graphics, etc.).

2. The parameters (all $w_i^{(k)}$) are **set to random values**



(3) The choice of a training method for the model



1. an image is chosen (at random) from the training set.
2. We know his label ("portrait" or not)

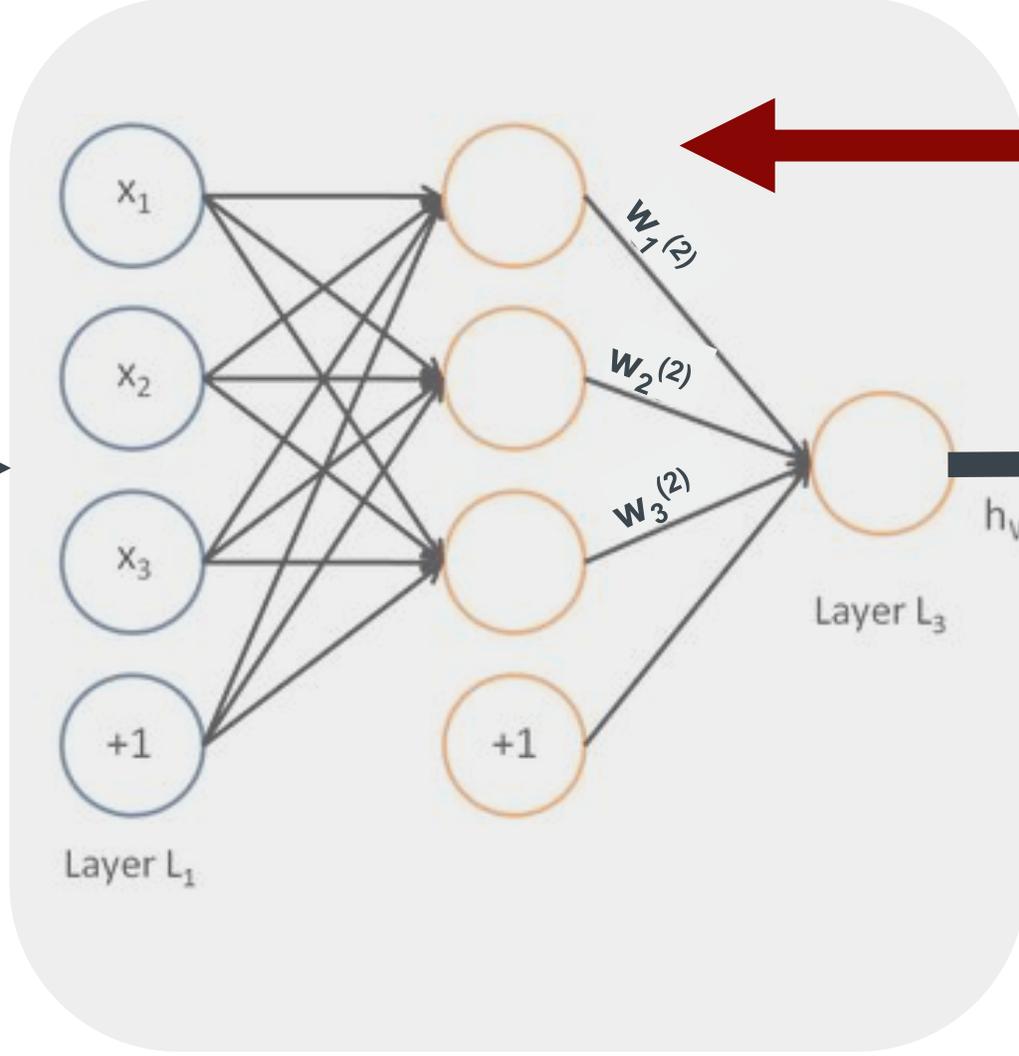
3. The pixels of an input image are supplied as input to the first layer of the network

Data training (known content)



1

image + label (portrait: "1")

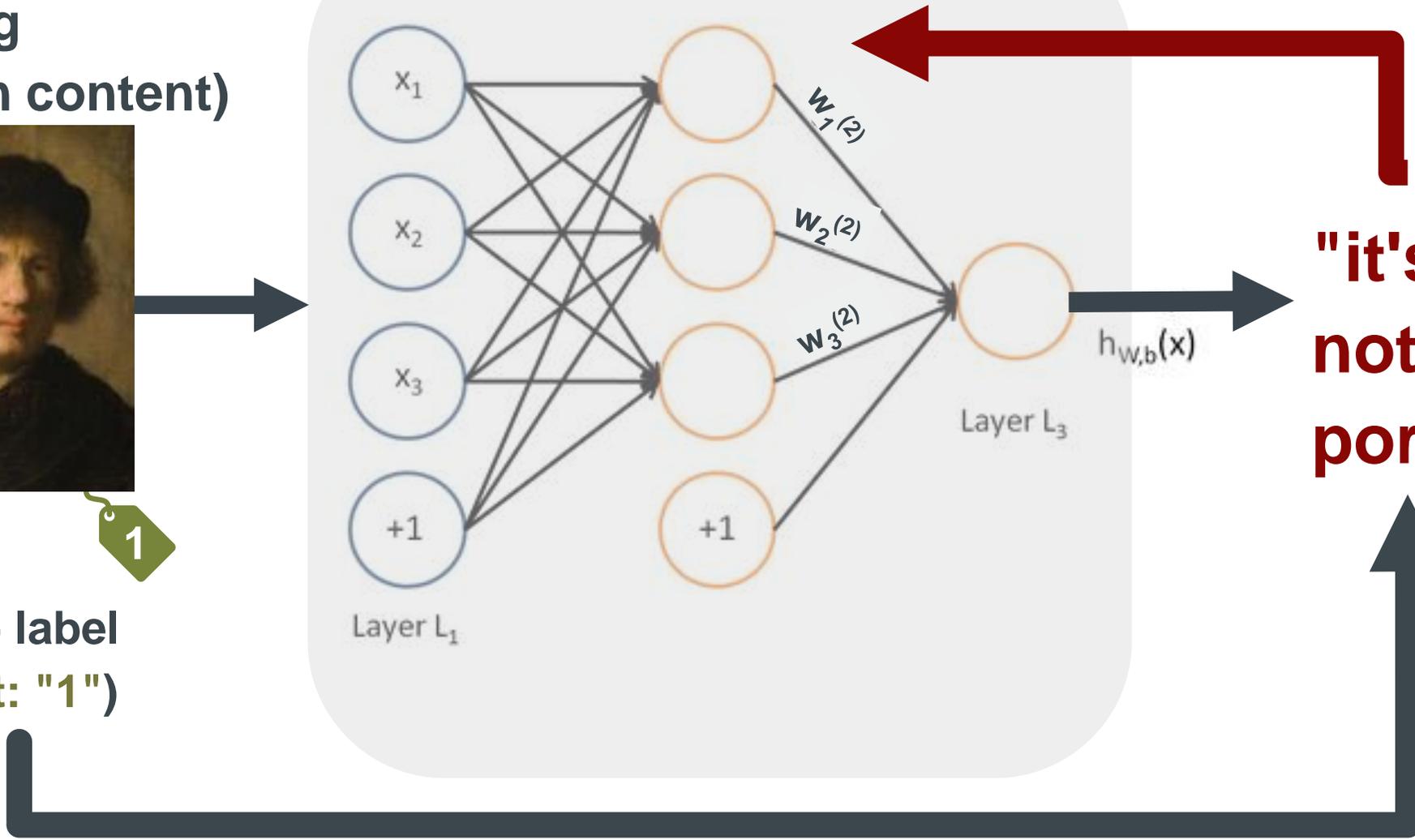


Modification of parameters

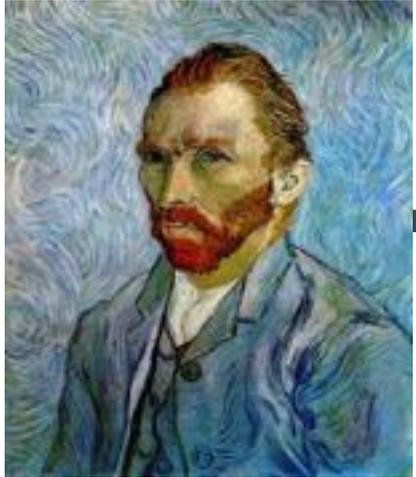
The weight of the neurons is modified according to their contribution to the wrong answer.

"it's not not a portrait" ~~X~~

we compare labels : "1" is expected

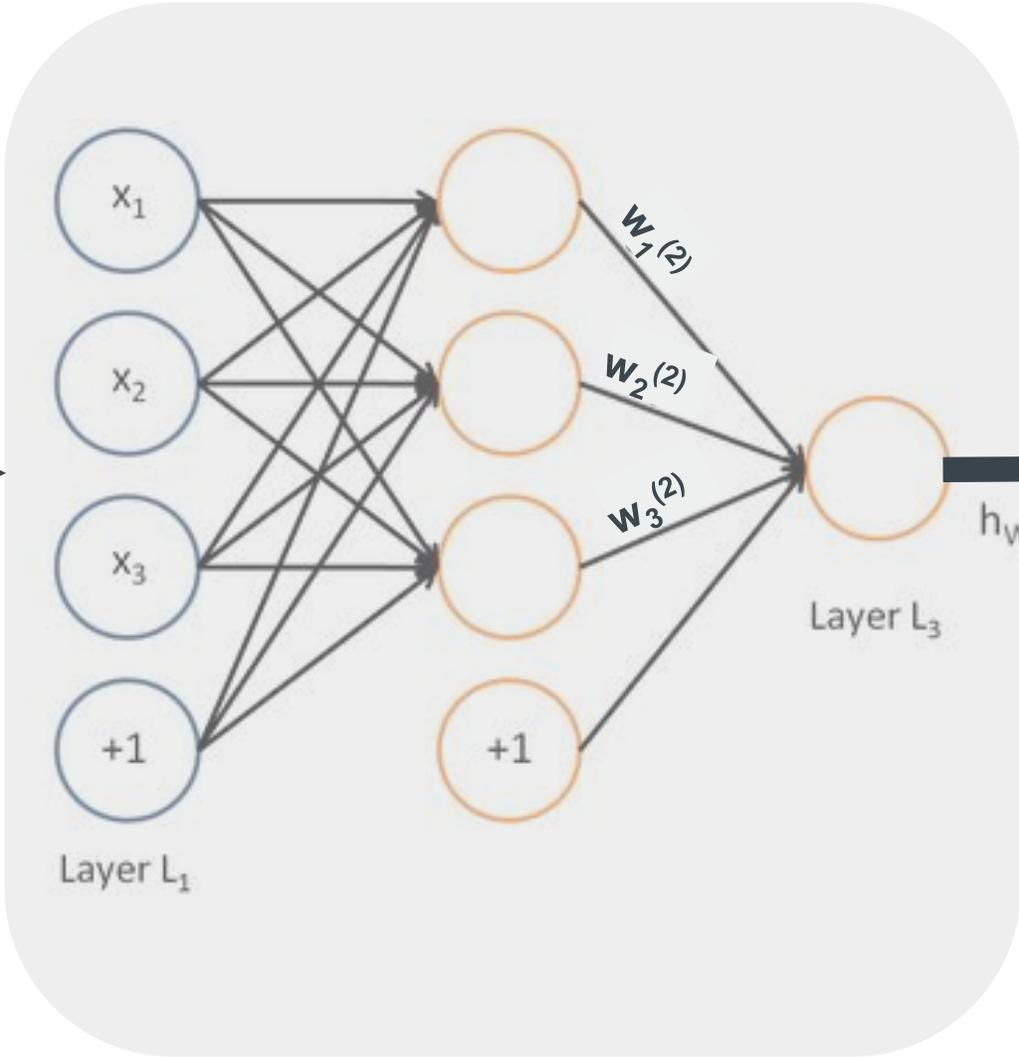


Data training (known content)



1

image + label (portrait)



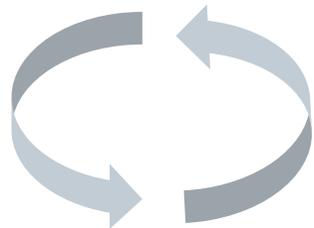
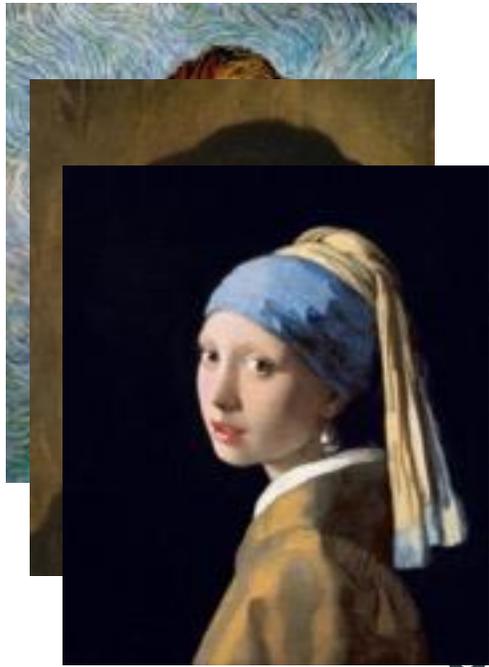
settings Ok



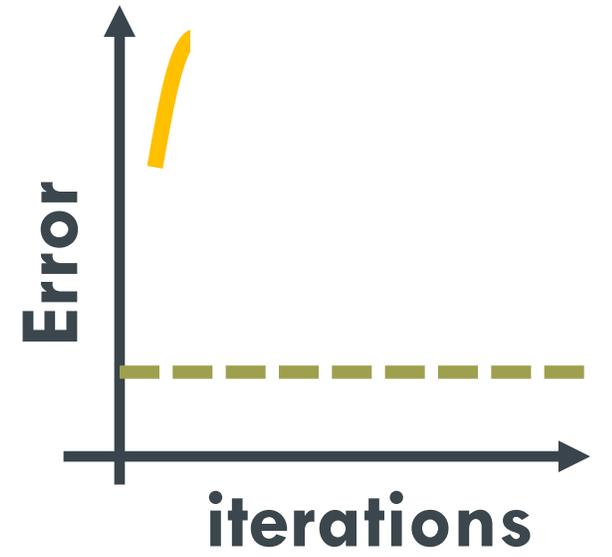
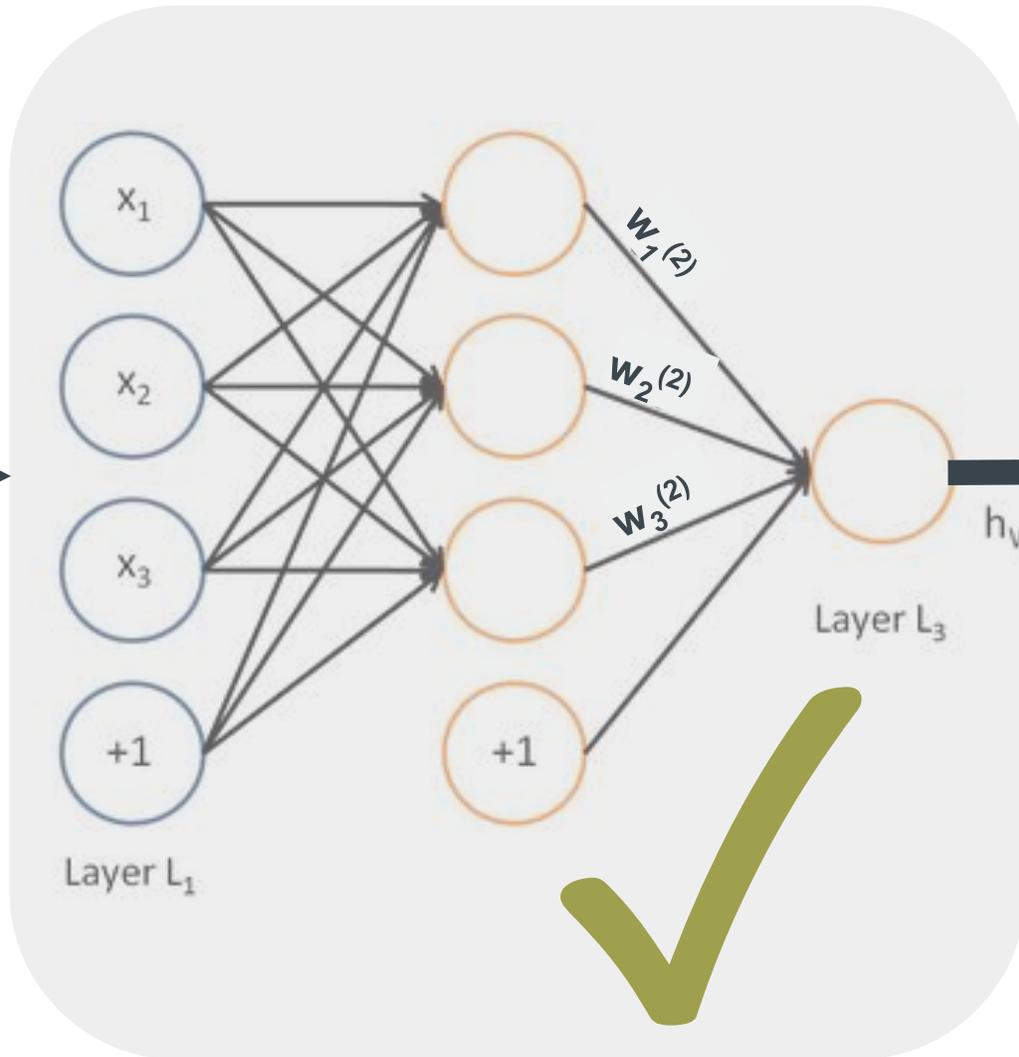
(we do nothing)

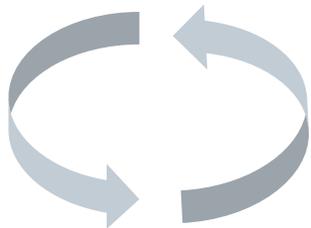
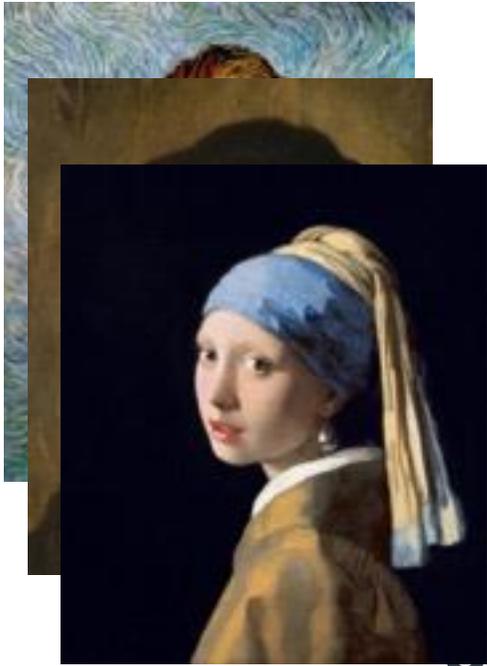
portrait ✓

The model is trained (the parameters are modified) until the "contract" is satisfied...

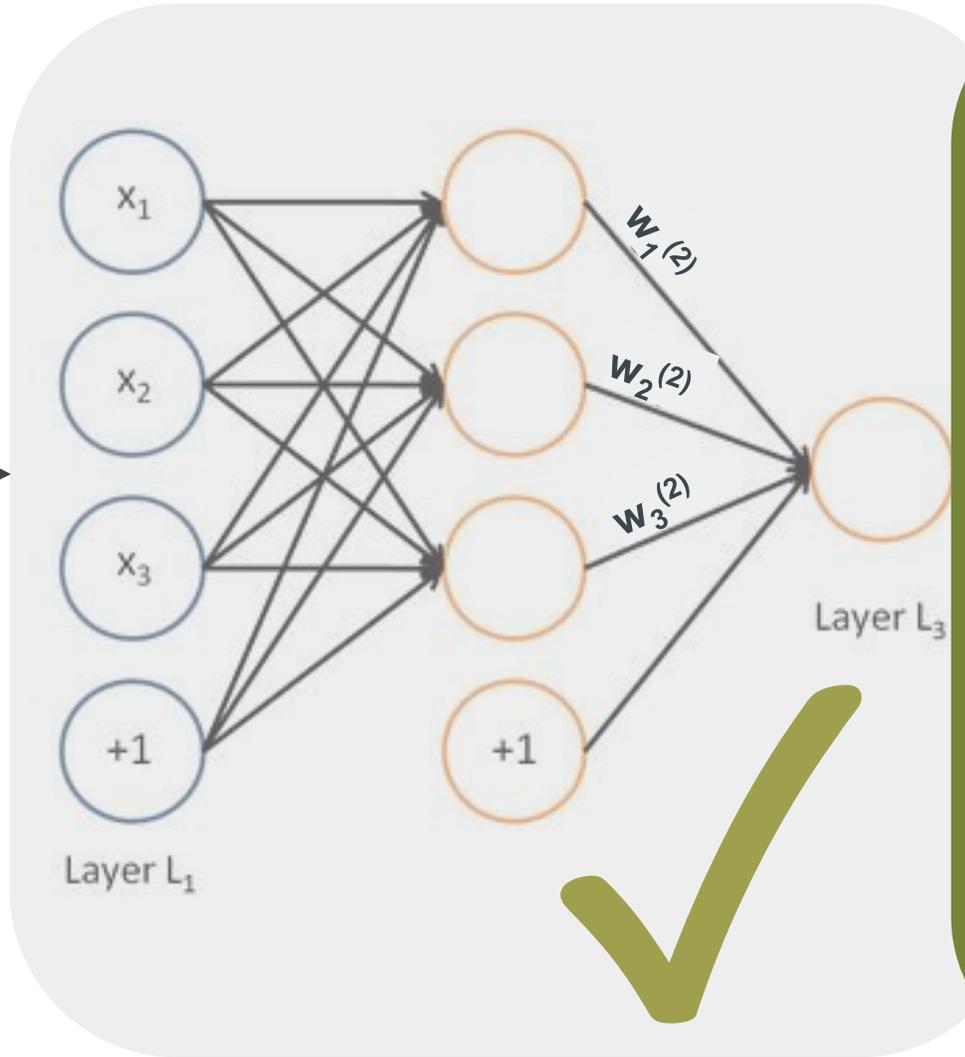


1





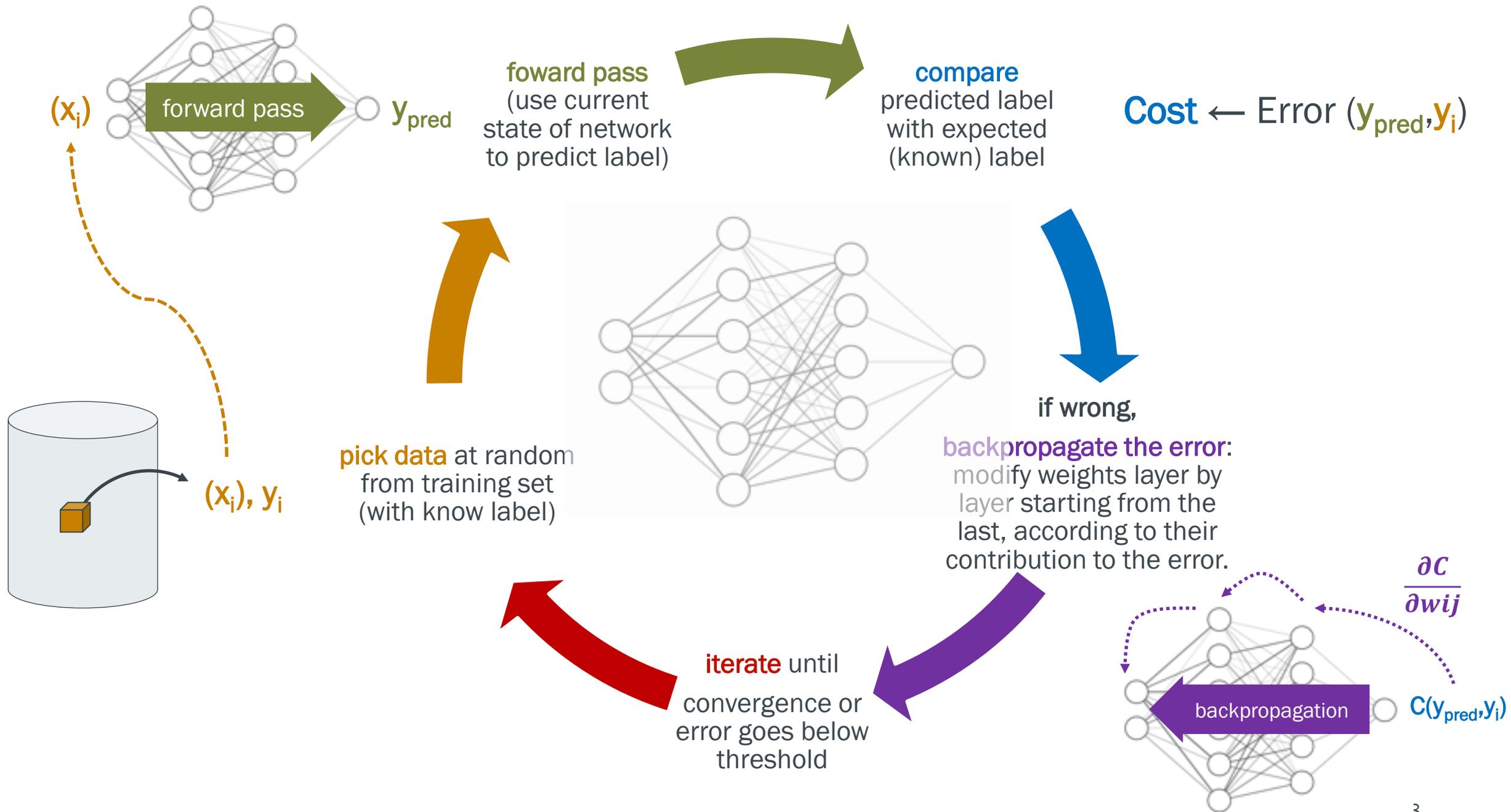
1



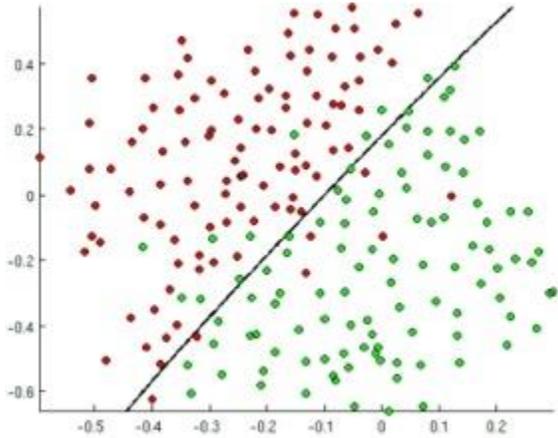
We used the training data to modify the parameters (the weights) of each neuron so that the overall behaviour of the network met a predefined quality metric (an error)

The result is a set of parameters ("weights") which control the operation of the NN.

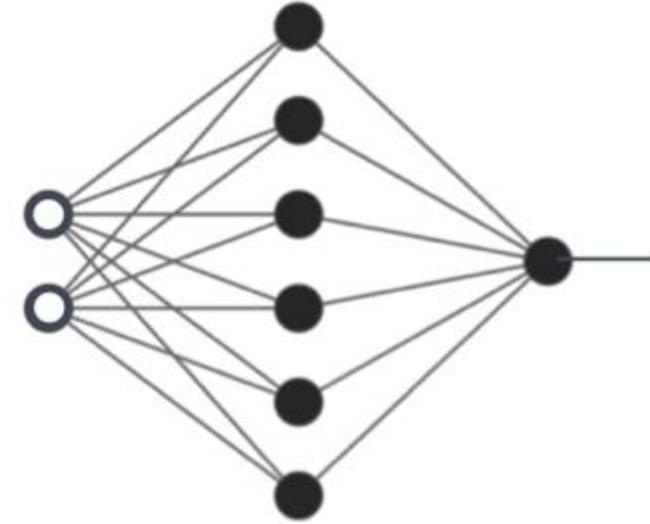
this trained model can now be applied to new data



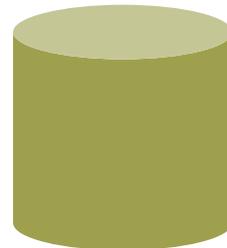
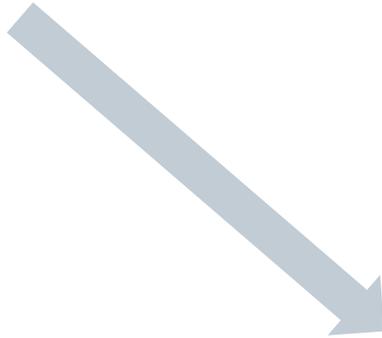
Link between complexity (of the problem) and model size



simple" model
(few parameters)

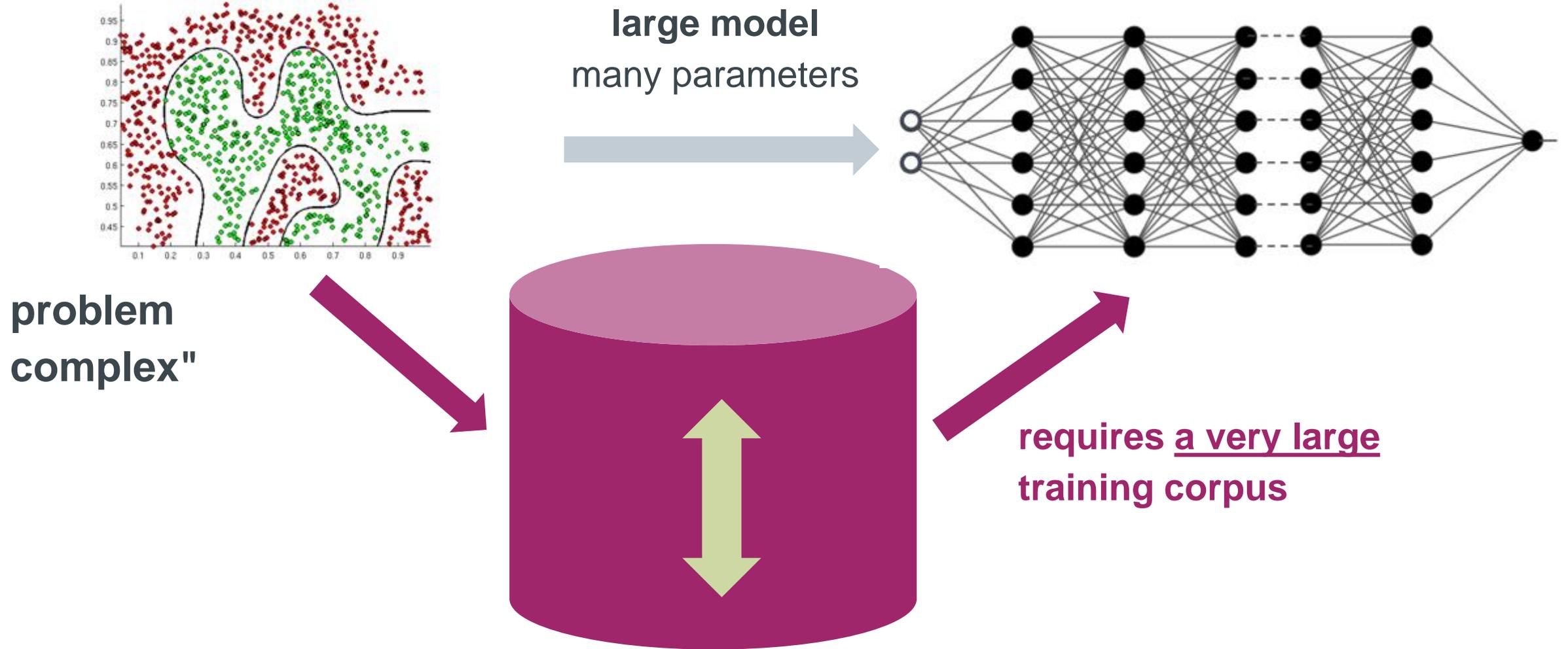


problem
"simple"

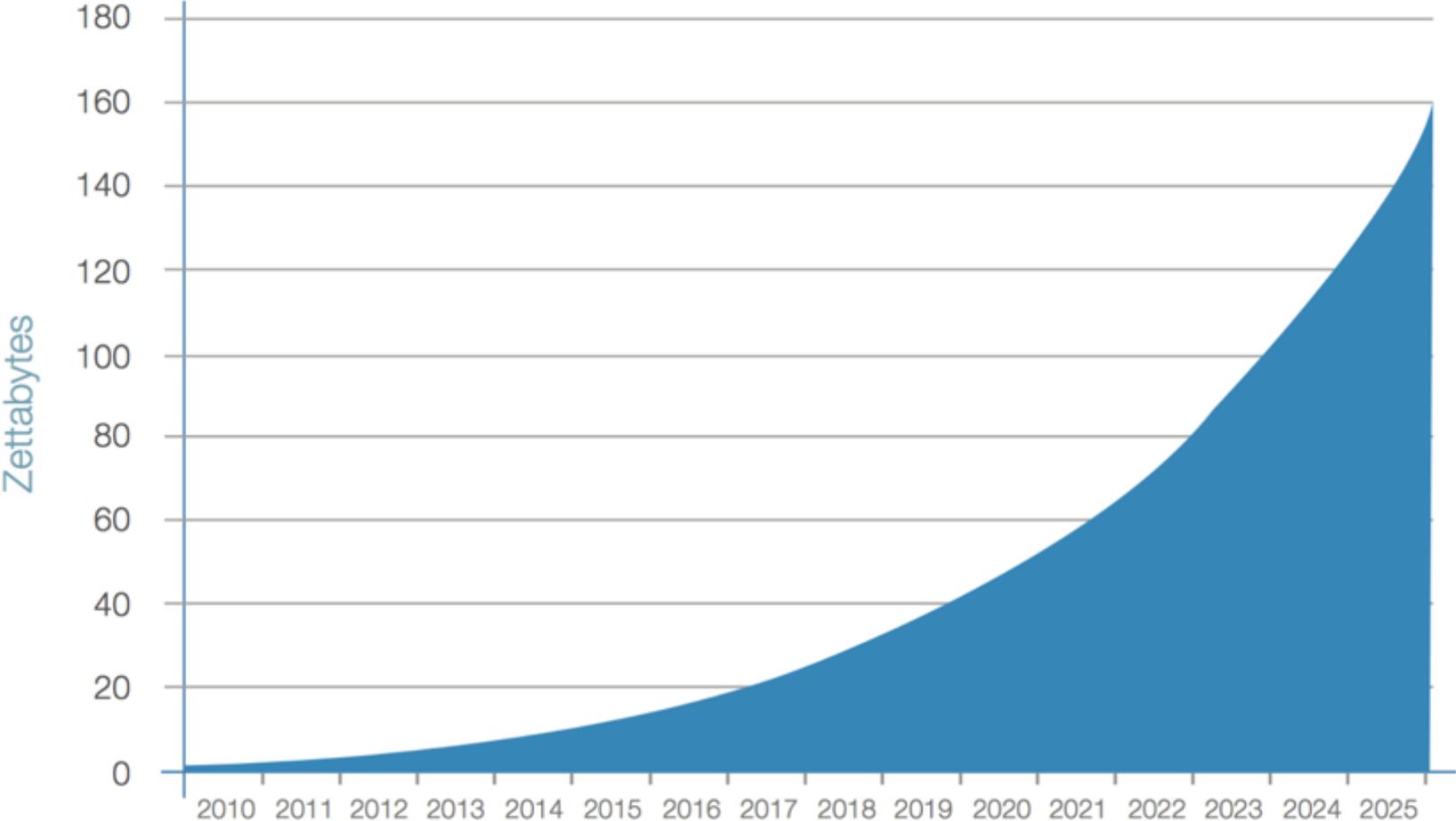


A small number of examples
training is all it takes

Link between complexity (of the problem) and model size



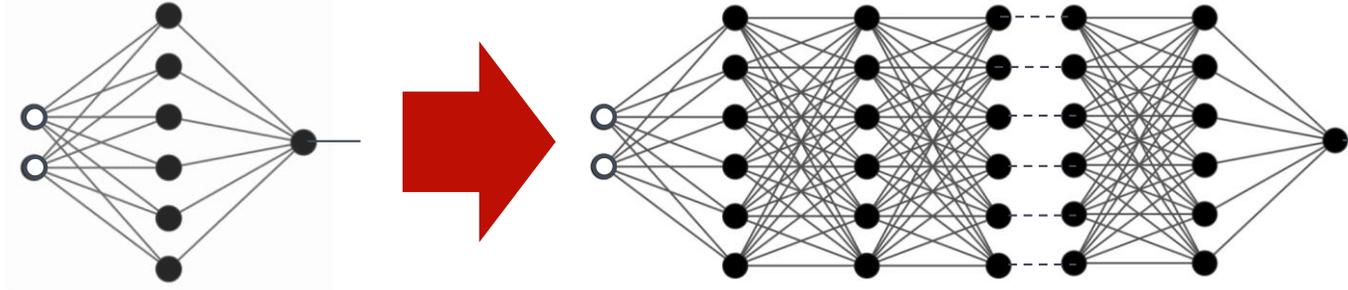
AI is becoming more and more data-hungry...



Source: IDC's Data Age 2025 study

Why the hype?

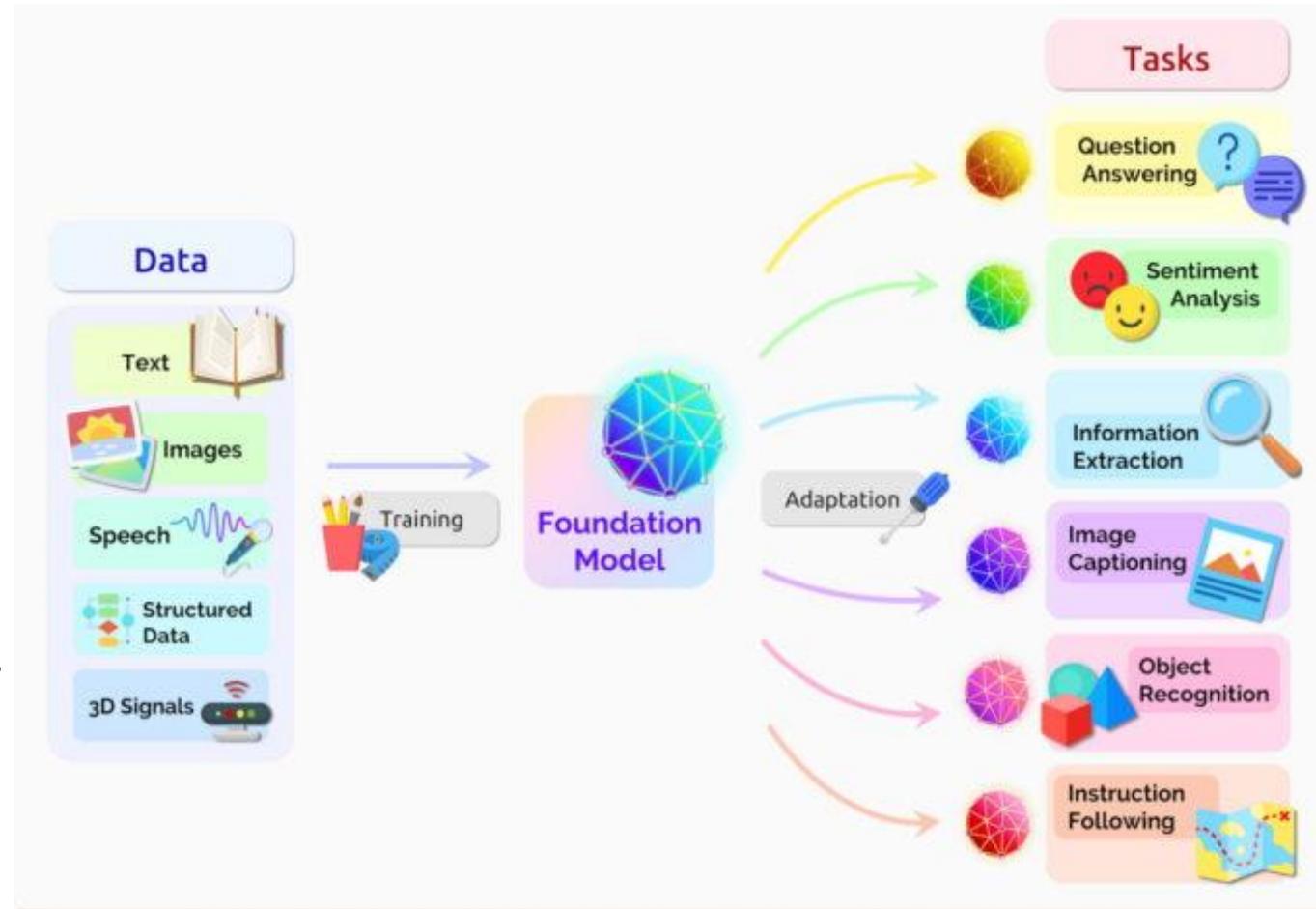
- Algorithmic progress
- Access to CPU (GPR/TPU...)
- Access to data



General Purpose AI / Foundation models

Foundation models /GPAI: large-scale AI models trained on vast amounts of data, designed to be adaptable to a wide range of downstream tasks. They serve as a base upon which various applications can be built, reducing the need for task-specific models.

- **Capable of performing multiple tasks** across different domains without task-specific training.
- **Trained on extensive datasets**, often using self-supervised learning techniques.
- **Can be fine-tuned for specific applications**, enhancing their versatility.



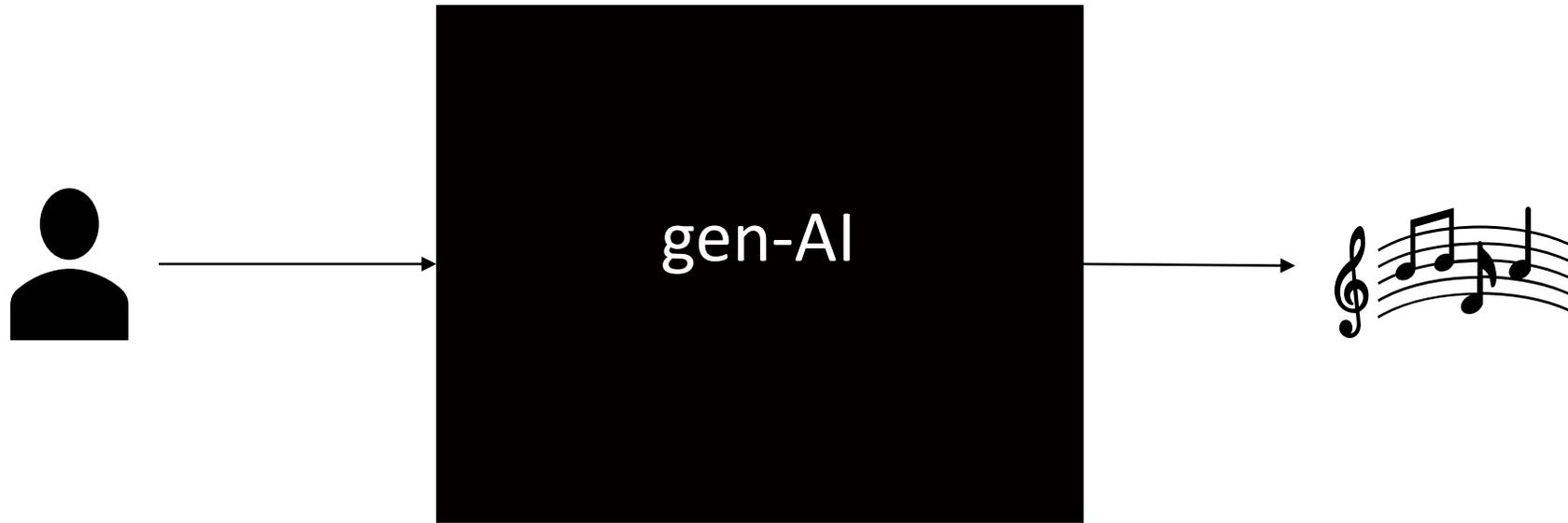
What does it mean “to learn”?

In AI, learning involves a system improving its performance on specific tasks by optimizing its internal parameters based on training data.

Humans define the model’s architecture (its structure and properties), choose the task to be solved, set the performance metrics, collect the training data, validate the outputs, etc.

What is generative AI?

How can algorithms produce novel artefacts?



Generative AI refers to algorithms that can create new content—such as text, images, audio, or video—by learning patterns from existing data.

How It Works:

- **Training Phase:** The model is exposed to vast amounts of data, allowing it to learn statistical patterns and relationships.
- **Generation Phase:** Upon receiving an input (often called a "prompt"), the model produces content by predicting and assembling elements based on the learned patterns.

Not so new (again)





• FAKE

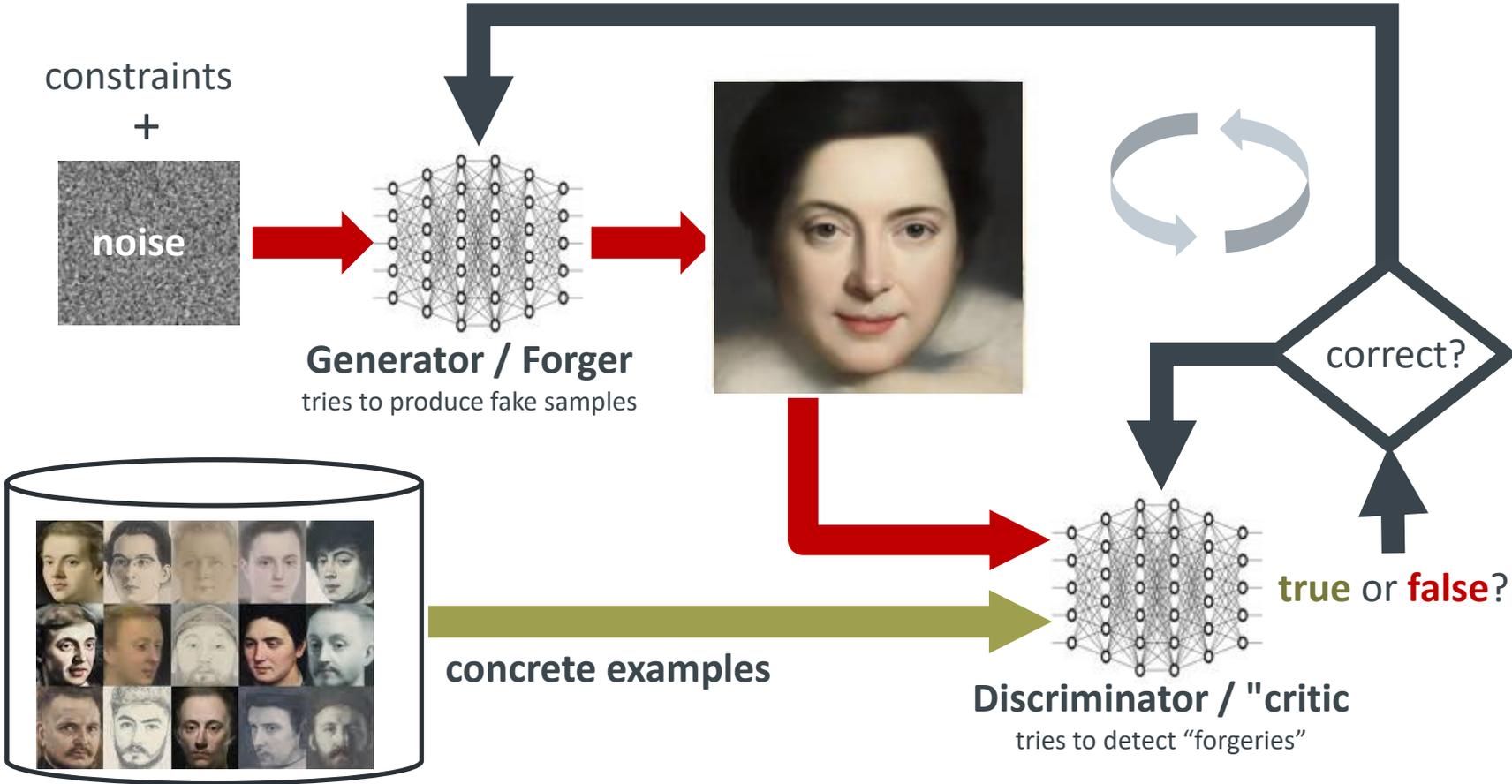


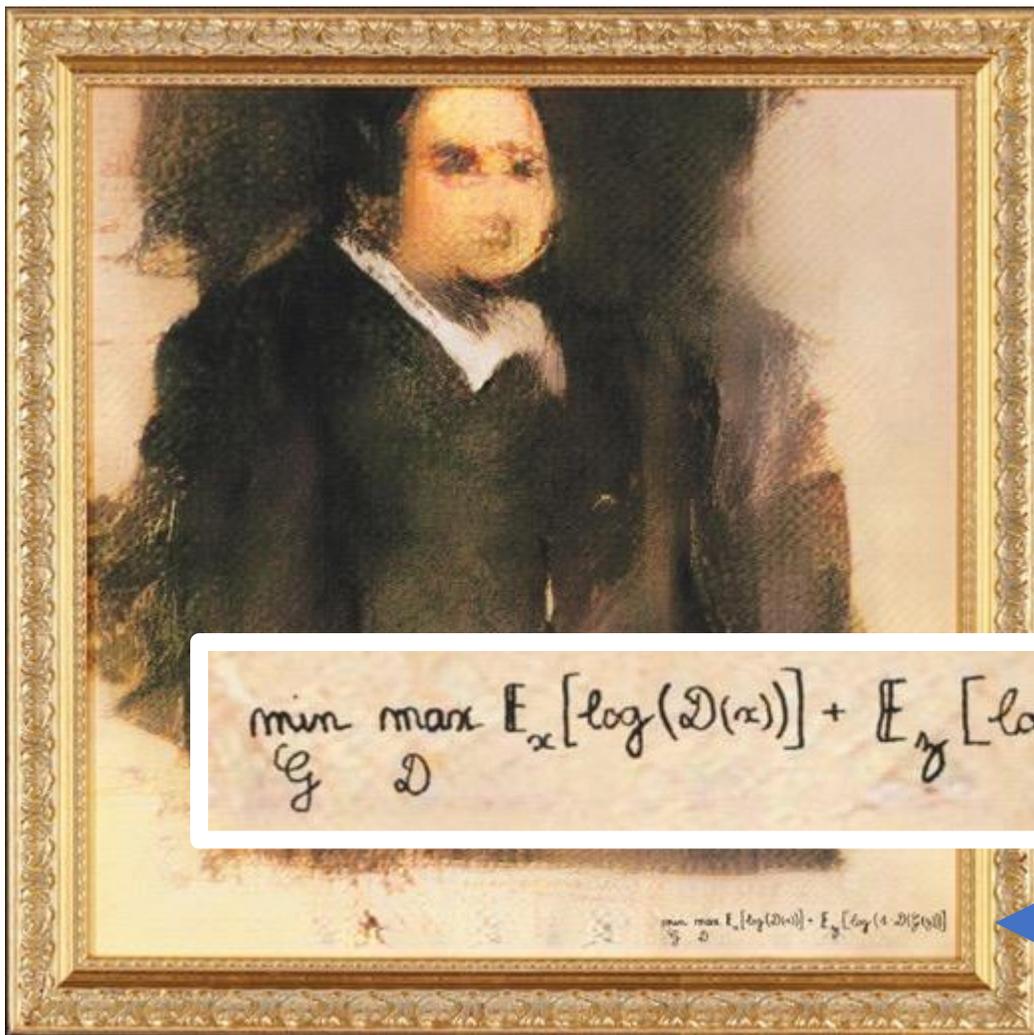
Style transfer



portraitAI.app

Generative AI & generative adversarial networks





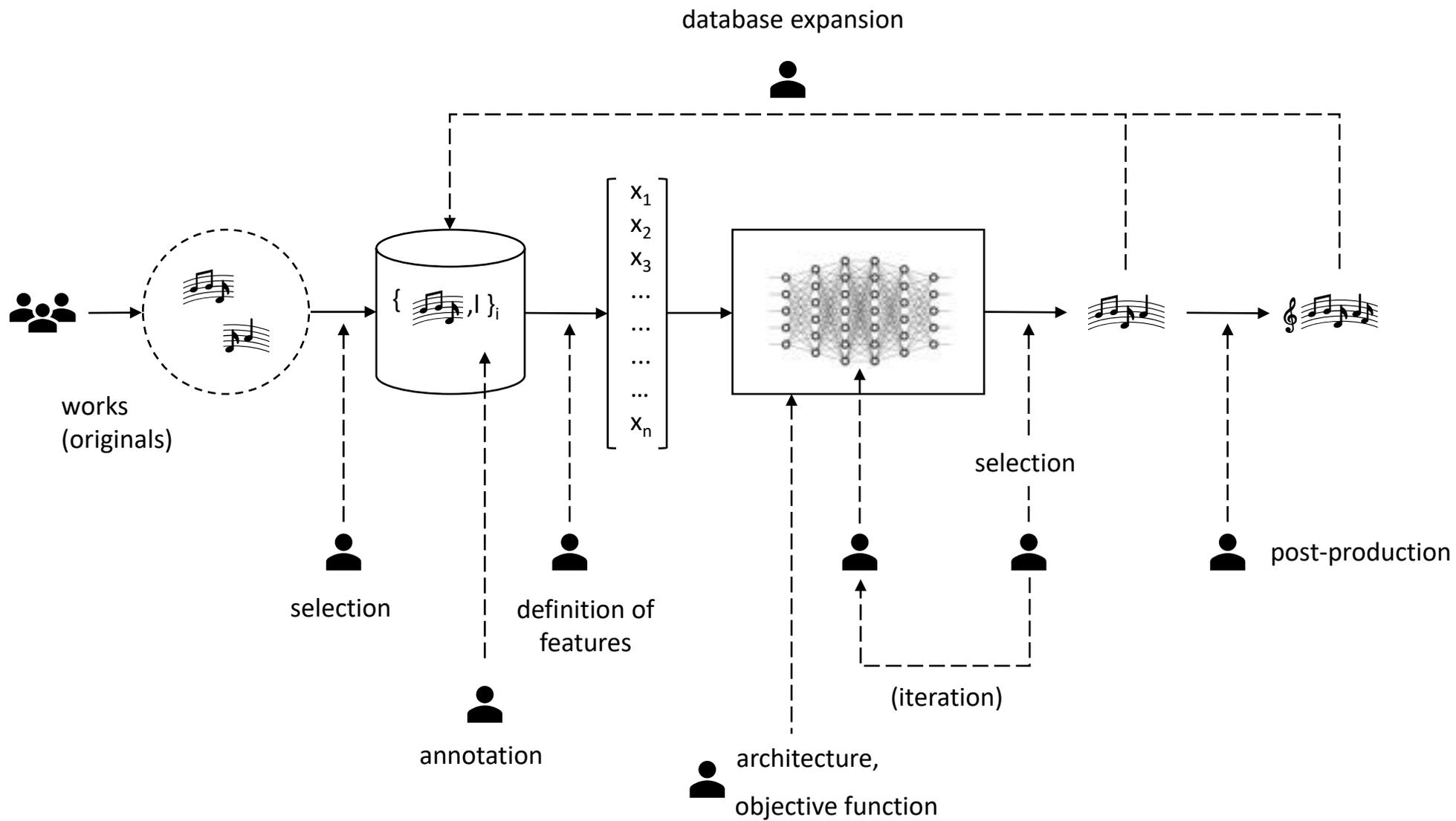
$$\min_G \max_D \mathbb{E}_x [\log(D(x))] + \mathbb{E}_y [\log(1 - D(G(y)))]$$

$$\min_G \max_D \mathbb{E}_x [\log(D(x))] + \mathbb{E}_y [\log(1 - D(G(y)))]$$

Today → Transformers, Diffusion models

1. progressively add noise to an image





What is generative AI?

Generative AI enables machines to create new content by learning patterns from existing data.

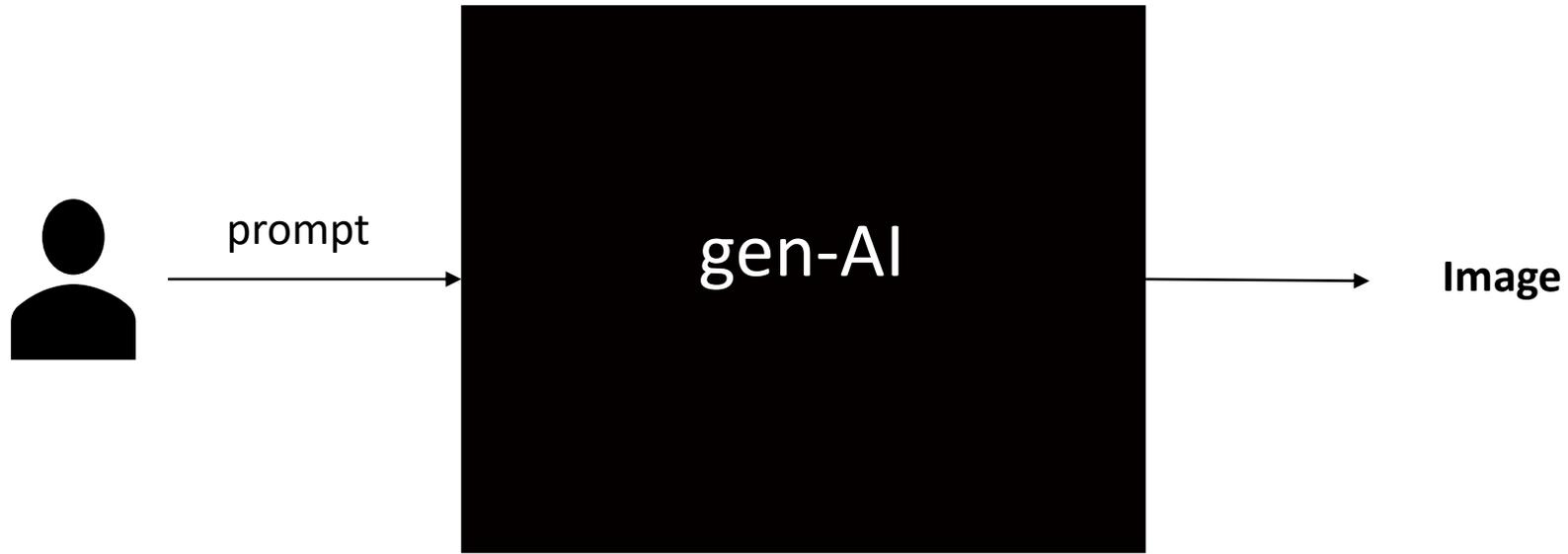
Who controls the outputs?

How much can we constrain the products coming out of generative AI?

Role of Prompts: The input prompt significantly influences the output. Carefully crafted prompts can guide the model toward desired responses, **but absolute control over the exact output is challenging due to the model's inherent complexity and probabilistic nature**. The model may produce unexpected results, especially with ambiguous or open-ended prompts.

Techniques for Output Control:

- **Prompt Engineering:** Designing specific prompts to steer the model's responses in a particular direction.
- **Fine-Tuning:** Adjusting the model's parameters on a specialized dataset to align outputs with specific requirements.
- **Sampling Strategies:** Methods like temperature setting and top-k sampling influence the randomness and diversity of the generated content.



What if we repeat (using the same prompt)? → same image?

What if we use more detailed prompts → more control over the output?

Basic: A cat.



chatGPT

Very Detailed: A peaceful domestic scene captured as if painted by a soft realist artist. A fluffy orange tabby cat with white paws lies curled on an aged oak windowsill, the morning sunlight filtering through sheer lace curtains, casting a warm glow on its fur. A small ceramic pot holding a vibrant green succulent sits beside the cat, creating a balance of color and texture. Outside the window, a blurred garden with blooming daisies and butterflies hints at spring. The cat looks serene, eyes half-closed, ears twitching faintly. The mood is calm, quiet, and filled with gentle warmth, evoking a sense of home and contentment. The artist's intent is to depict a fleeting, tender moment of stillness in a chaotic world.

- **What if more words, more details?...**

chatGPT



Bing



Who controls the outputs?

While prompts and tuning techniques can guide generative AI outputs, achieving control over the content and form is inherently limited: the details of the expression are forged by the model (its architecture, training data, fine-tuning, human-feedback @ training, validation)

Do models “memorize”?

Gen-AI as a “stochastic parrot”?

Model memorization (of training data)

Memorization vs. Generalization:

Memorization occurs when a model recalls exact details from its training data, whereas generalization refers to the model's ability to apply learned patterns to new, unseen data.

Why?

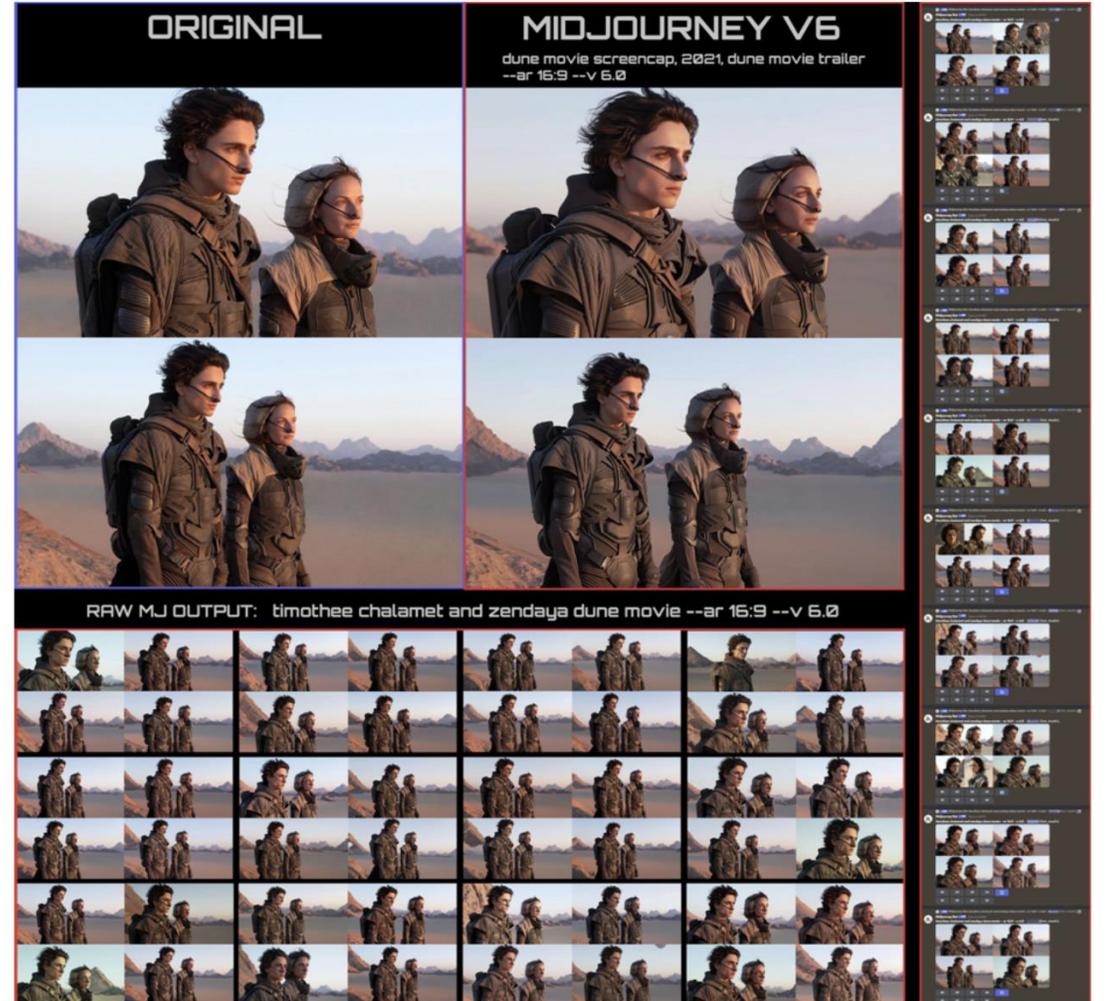
- **Model Size:** Larger models with more parameters have a higher capacity to memorize training data.
- **Training Duration:** Extended training periods can lead to overfitting, where the model becomes overly specialized to the training data.
- **Data Characteristics:** Imbalanced or small datasets increase the likelihood of memorization.

MIDJOURNEY (AND CO)

FILM FRAME

MIDJOURNEY V6

Thanos infinity war, 2018, screenshot from a movie,
movie scene, 4k, bluray --ar 16:9 --v 6.0



(large) generative models as memory banks...

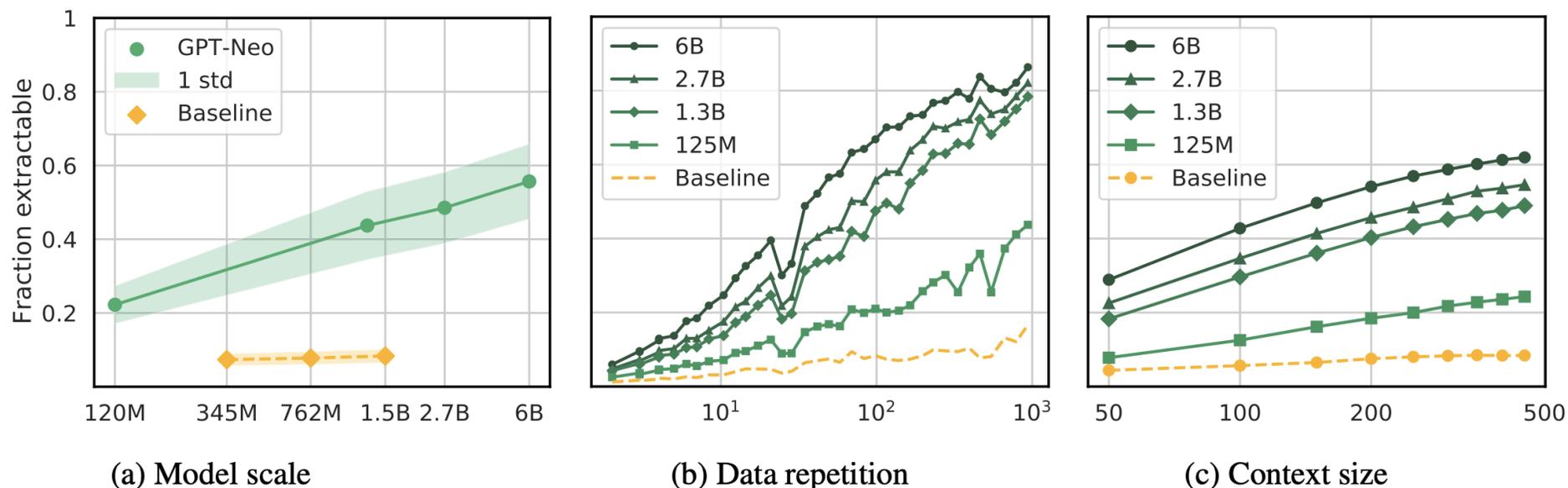


Figure 1: We prompt various sizes of GPT-Neo models (green) with data from their training set—The Pile. As a baseline (yellow), we also prompt the GPT-2 family of models with the same Pile-derived prompts, even though they were trained on WebText, a different dataset. **(a)** Larger models memorize a larger fraction of their training dataset, following a log-linear relationship. This is not just a result of better generalization, as shown by the lack of growth for the GPT-2 baseline models. **(b)** Examples that are repeated more often in the training set are more likely to be extractable, again following a log-linear trend (baseline is GPT-2 XL). **(c)** As the number of tokens of context available increases, so does our ability to extract memorized text.

Do models “memorize”?

Large AI models can memorize a small fraction of the training data, which may lead to overfitting and potential reproduction of specific content.

Mitigation strategies : work in progress!

Thank you!

ARENAIRE

INTELLECTUAL PROPERTY

AI AS A SOURCE OF BOTH CREATION AND INFRINGEMENT

Part 1: Protection of output generated by AI: what say the IP Offices?

Pierre Massot

**3 april
2025**



Foreword

Much ado about nothing?

Business as usual?

What for tomorrow?



Created by MidJourney with the prompt « Much
ado about nothing »

Foreword

For the time being, lawyers have 2 practical questions:

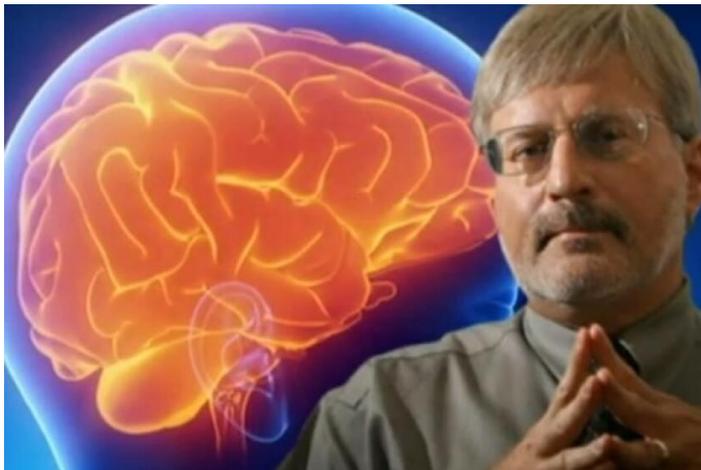
- Is it possible to obtain IPRs when a creation or an invention is solely created or invented by an AI system?
- Is it possible to obtain IPRs when a human being creates or invents with the help of AI? If yes, at what conditions?

1. Creativity of AI (only)

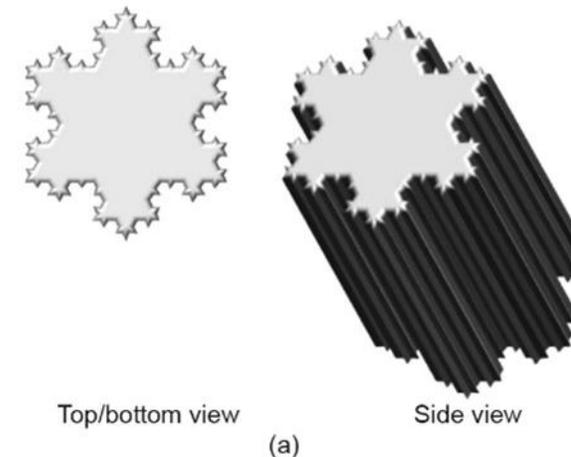
The (quite unsuccessful) fight of Stephen Thaler & the Artificial Inventor Project

“a small but growing number, are fighting a decidedly uphill battle in challenging the deep-seated belief that only a human can invent”.

Steve Lohr, 15 July 2023, Can AI invent? NYT, www.nytimes.com/2023/07/15/



Source: www.managingip.com



Inventions designed by the AI system DABUS and included in a patent application (Food container whose walls have a fractal profile, allowing it to fit together with others and improving grip and heat transfer into and out of the container)

1.1. Inventions generated by AI only

No inventorship for AI (usually)

Relative global consensus: AI cannot be an inventor

- **USA: Inventors named on U.S. Patent Applications must be natural persons** (Guidance on Use of AI-Based Tools, USPTO)
- **China: Inventor named in the patent document must be a natural person.** AI systems and other non-natural persons cannot be inventors (Guidelines for Patent Application, CNIPA, 31/12/24)
- **EPO: Every application must designate the inventor, who must be a natural person** (EPO Guidelines + EPO, Board of Appeal, case J8/20, 21/12/2021, DABUS*)

Exception: South Africa (but no formal examination by the office)

*DABUS (Device for the Autonomous Bootstrapping of Unified Sentience) is an AI system created by Stephen Thaler

1.2. Creations generated by AI (only)

No authorship for AI (usually)

➤ Same global trend for copyright

Ex1: US Copyright Office, Copyright and Artificial Intelligence, Part 2, January 2025:

“copyright protection in the United States requires human authorship. This foundational principle is based on the Copyright Clause in the Constitution and the language of the Copyright Act as interpreted by the courts. The Copyright Clause grants Congress the authority to “secur[e] for limited times to authors... The exclusive right to their... writings.”³³ As the Supreme Court has explained, “the author [of a copyrighted work] is... the person who translates an idea into a fixed, tangible expression entitled to copyright protection”

No court has recognized copyright in material created by non-humans, and those that have spoken on this issue have rejected the possibility”

Ex2: EU law (an intellectual creation is an author’s own if it reflects the author’s personality ; ECJ, C-145/10, Painer)

2. Human creativity assisted with AI

2.1. Inventions made by humans with the help of AI

AI-assisted inventions: a possible protection

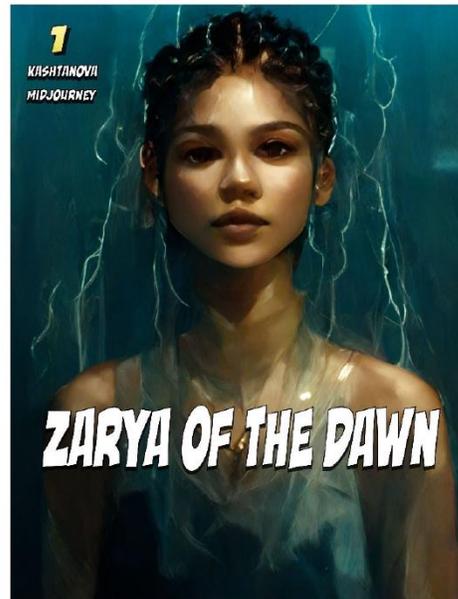
- **USA:** The use of an AI system by a natural person does not preclude a natural person from qualifying as an inventor if the natural person significantly contributed to the claimed invention (Guidance on Use of AI-Based Tools, USPTO)
- **China:** For inventions assisted by AI, a natural person who has made a creative contribution to the substantive features of the invention can be named as the inventor of the patent application (Guidelines for Patent Application, CNIPA, 31/12/24)

2.2. Creations made by humans with the help of AI

The (still) restrictive approach of the US Copyright office

US Copyright Office, February 21, 2023

- Ms. Kashtanova is the author of the Work's **text** as well as the **selection, coordination, and arrangement** of the Work's written and visual elements. That authorship is protected by copyright.
- **But** : the images in the Work that were generated by the Midjourney technology are **not** the product of human authorship. No copyright.



Zarya of the Dawn²

Reasons of the decision

- Midjourney does not interpret prompts as specific instructions to create a particular expressive result
- The process is not controlled by the user because it is not possible to predict what Midjourney will create ahead of time
- Accordingly, Midjourney users are not the “authors” for copyright purposes of the images the technology generates



Case Jason M. Allen

- Mr. Allen “*input numerous revisions and text prompts at least 624 times to arrive at the initial version of the image*”
- He used **Adobe Photoshop to remove flaws** and create new visual content and used **Gigapixel AI to “upscale”** the image, increasing its resolution and size

Théâtre D’opéra Spatial,

An A.I.-Generated Picture Won
an Art Prize. Artists Aren’t
Happy, NYT, Sept. 2, 2022
<https://www.nytimes.com>



US Copyright Office, September 5, 2023

Refusal of the application because Mr. Allen is unwilling to disclaim the AI-generated material:

- **Mr. Allen's actions do not make him the author of the Midjourney Image. His sole contribution was inputting the text prompt.**
- Prompts “influence” what the system generates **BUT Midjourney does not interpret prompts as specific instructions to create a particular expressive result**
- If the process of prompting can involve creativity, **that does not mean that providing text prompts to Midjourney “actually form[s]” the generated images**

US Copyright Office, 30 January 2025

- Registration on the basis of the “selection, coordination, and arrangement of material generated by artificial intelligence”
- **But** exclusion of: « *2-D artwork, AI generated image components* »

A Single Piece of American Cheese

Visual material generated
by [Invoke AI](#) with a
technique called
“inpainting”



A broader approach elsewhere?

Ex.: Ukraine



- *Ukrainian IP Office issued a copyright certificate for the composite work – a collection of images “Easter cards”, where images generated by a computer program were used.*

* <https://ukrpatent.org/en/articles/UKRNOIVI-about>

Ex.: Japan

- **Materials autonomously generated by AI:** not "creatively produced expressions of thoughts or sentiments" / no copyrighted "works"
- **If AI is used as a "tool" by a person to creatively express thoughts or sentiments:** considered a "work" + the user of the AI is the "author"
- **Determining whether a person has used AI as a "tool" depends on two factors:** whether the person had a "creative intention" and whether the person has made a "creative contribution"

"General Understanding on AI and Copyright in Japan", Japan Copyright Office (JCO), Copyright Division, Agency for Cultural Affairs, Japan, May 2024 (not legally binding)

https://www.bunka.go.jp/english/policy/copyright/pdf/94055801_01.pdf

Ex.: Japan

Examples of factor in determining copyrightability of AI-generated material

- ✓ Amount of instructions/input: “Detailed instructions that specifically indicate what constitutes creative expression” are more likely to be considered as creative contributions. However, lengthy instructions (i.e., prompts) that merely suggest an idea do not influence the assessment of creative contribution.
 - ✓ Number of generation attempts: A large number of attempts alone does not affect the assessment of creative contribution. Repeated attempts, while checking the generated materials and correcting the instructions/input, may be recognized as a creative contribution.
 - ✓ Selection from multiple output materials: The mere act of selection itself does not influence the determination of creative contribution. However, certain elements of choice may be involved which may be considered as creative.
- In addition, any additions or corrections made by humans to AI-generated materials that can be considered creative expressions are generally considered to be copyrighted works.

“General Understanding on AI and Copyright in Japan”, Japan Copyright Office (JCO), Copyright Division, Agency for Cultural Affairs, Japan, May 2024 (not legally binding)

CONCLUSION

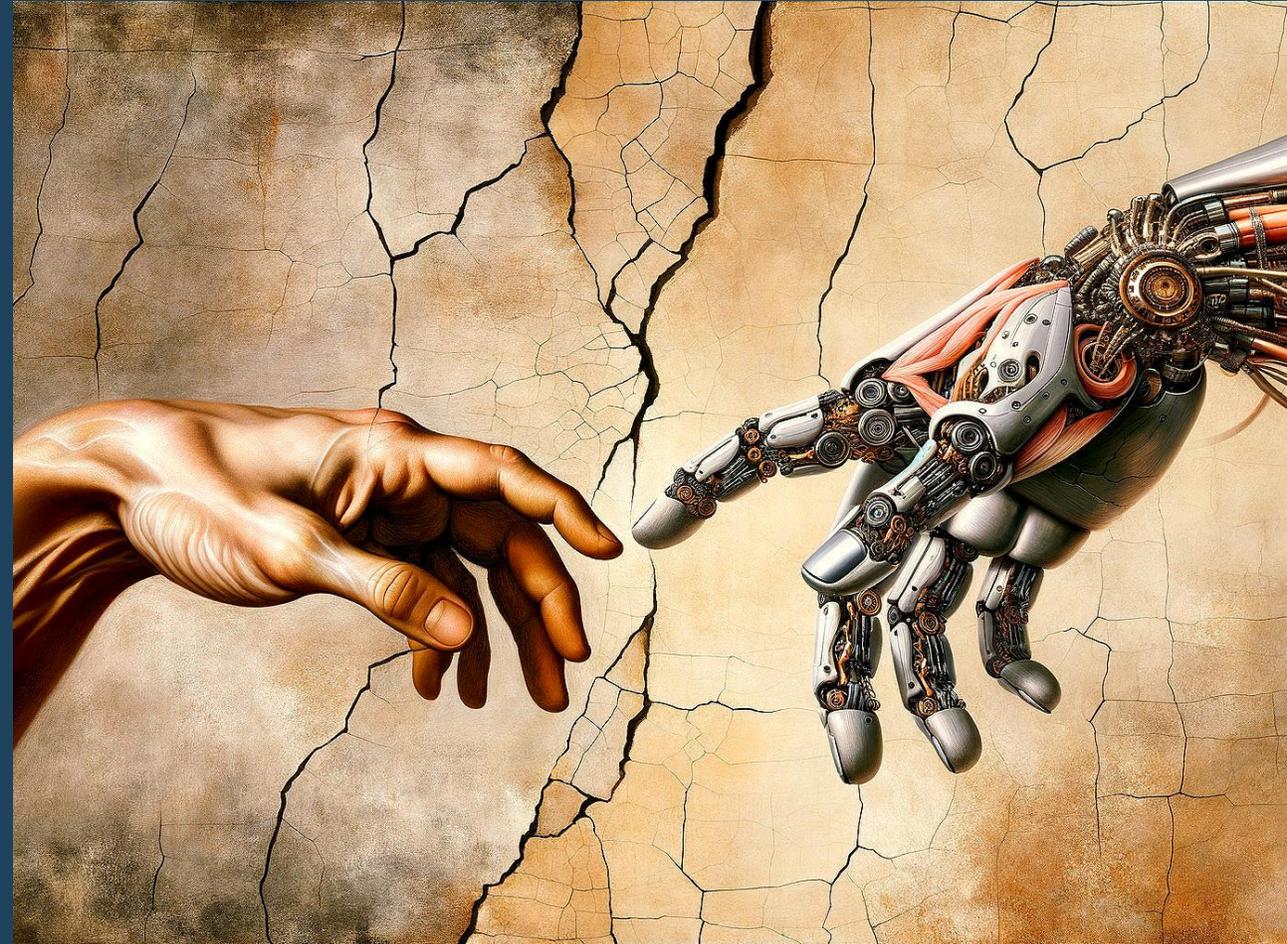
- The global trend is to refuse IPR's for inventions and creations generated only by AI
- Inventions and Creations generated with the assistance of AI could / should be protectable if there is some human creativity

Thank you for your attention !



AI as a source of creation: An international perspective

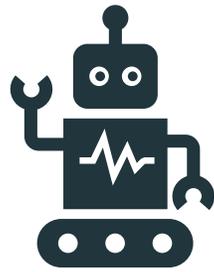
PART ONE



USA: Thaler v. Perlmutter

1. *U.S. District Court for the District of Columbia, Thaler v. Perlmutter, Civil Action No. 22-1564 (BAH), August 18, 2023*
2. *US Court of Appeals, District of Columbia Circuit, Thaler v. Perlmutter, Civil Action No. 23-5233, March 18, 2025*

1. U.S. District Court for the District of Columbia, Thaler v. Perlmutter, Civil Action No. 22-1564 (BAH), August 18, 2023



1

Stephen Thaler created an AI system called the “Creativity Machine.”



2

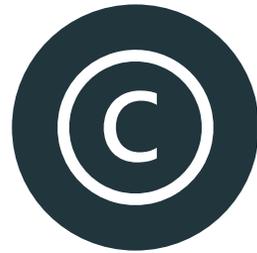
This AI generated a visual artwork titled “A Recent Entrance to Paradise.”

A Recent Entrance to Paradise generated by Creativity Machine



Image Thaler v. Perlmutter decision

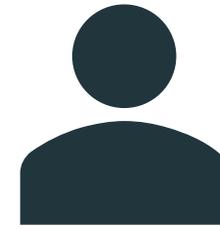
1. U.S. District Court for the District of Columbia, Thaler v. Perlmutter, Civil Action No. 22-1564 (BAH), August 18, 2023



3

Thaler applied to register the artwork for copyright.

In his application, he listed the **AI as the author** and himself as the copyright owner (because he owns the machine).



4

The U.S. Copyright Office denied the request, saying that **only humans can be authors** under US copyright law.

1. U.S. District Court for the District of Columbia, *Thaler v. Perlmutter*, Civil Action No. 22-1564 (BAH), August 18, 2023



Question

Can a work created entirely by AI, with no human involvement, be protected by U.S. copyright law?

1. U.S. District Court for the District of Columbia, Thaler v. Perlmutter, Civil Action No. 22-1564 (BAH), August 18, 2023



Short answer

No.

1. U.S. District Court for the District of Columbia, *Thaler v. Perlmutter*, Civil Action No. 22-1564 (BAH), August 18, 2023



Reasoning

The US Copyright law protects **only works created by human beings**.

“The plaintiff did correctly observe that throughout its history, copyright law has proven malleable enough to cover works created with or involved technologies developed long after traditional media of writings memorialized on paper”.

It has been consistent *“that human creativity is sine qua non at the core of copyrightability, even as that human creativity is channelled through new tools or into new media”.*



Older cases

Burrow-Giles v. Sarony (1884): Photographs copyrightable due to human creative choices.

Naruto v. Slater (2018): Animals cannot hold copyright.



Ruling

In the **absence of human involvement, the work was never eligible for copyright.**

→ The Copyright Office acted properly in denying registration.

2. US Court of Appeals, District of Columbia Circuit, Thaler v. Perlmutter, Civil Action No. 23-5233, March 18, 2025



Appeal

The US appeals court upheld the decision by the lower court.



USCO's report

USCO's report Copyright and Artificial Intelligence Part 2: Copyrightability: *"if content is entirely generated by AI, it cannot be protected by copyright"*.



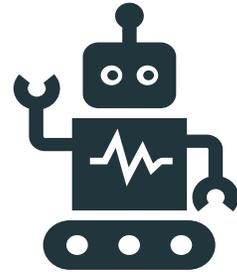
Human

"The human authorship requirement does not prohibit copyrighting work that was made by or with the assistance of AI. The rule requires only that the author of that work be a human being-the person who created, operated, or used AI-and not the machine itself".

China: Feilin v. Baidu

*Beijing Internet Court, 25 April 2019, No 239
[2019], Civil First Instance, Beijing Feilin Law
Firm v Baidu Corporation*

Feilin v. Baidu, April 25, 2019



1

The report "*Analysis Report on Judicial Big Data of the Entertainment Industry-Film Volume in Beijing*" generated by the software Wolters Kluwer China Law & Reference is published on a law firm's WeChat.

A user published the report without permission on the platform Baidu operated by the defendant.



2

Plaintiff:

Copyright infringement claims against Baidu.

Defendant:

No protection because the entire report was generated by the platform.

Feilin v. Baidu, April 25, 2019

Condition of protection

Graphic works

For the same data, different users would obtain the same graphs if applying the same software for processing; even if different software products are used, the expression of the same data should be the same as long as users display data in conventional graph types.

Written works

- created by a natural person within the scope of literature, art and science
- written form
- can be reproduced
- original

In this case:

After two tests, the Plaintiff's report is original, because a natural person added his own elements.

Feilin v. Baidu, April 25, 2019



Natural person

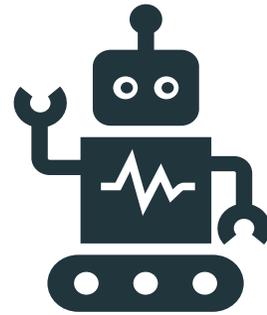
Originality is not sufficient on its own, the work must be created by a **natural person**.

- The software developer does not create the report → not the author
- The software user chooses keywords, but the report is generated automatically by the “visualisation function” → not the author without additional elements

China: Lin Chen v. Hangzhou Gaosi Membrane Technology

*Changshu Court in Jiangsu Province,
October 18, 2024*

Lin Chen v. Hangzhou Gaosi Membrane Technology, October 18, 2024



1

Lin Chen generated an image using AI tools MidJourney and Photoshop.

Publication on XiaoHongShu: February 14, 2023

Registration with China's Copyright Administration as a 2D artwork: April 7, 2023

Lin Chen v. Hangzhou Gaosi Membrane Technology, October 18, 2024



With Heart

Image from iam-media.com

Lin Chen v. Hangzhou Gaosi Membrane Technology, October 18, 2024

Conditions of protection

Although the picture was generated by an AI software, Chen had made changes that reflected his selection and arrangements of scene, environment, colour, light, shade, and angle of various elements in the picture.

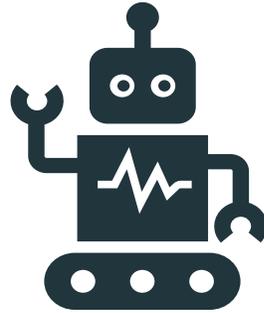
→ Original



China: LI v. LIU

Beijing Internet Court Civil Judgment (2023)
Jing 0491 Min Chu No. 11279

LI v. LIU, November 27, 2023



LI used Stable Diffusion AI software to generate a picture titled “*Spring Breeze Brings Tenderness.*”

He posted it on social media platform Little Red Book with his watermark.

“Spring Breeze Brings Tenderness” creation process



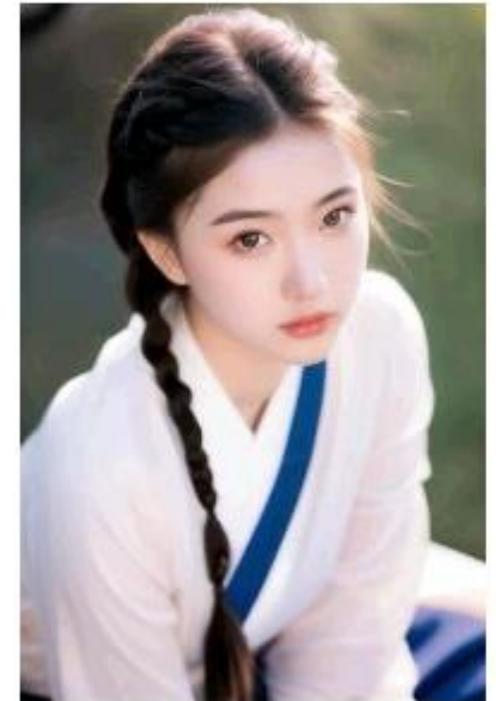
Step 6

Step 7



Step 8

Step 9



Images from LI v. LIU decision

LI v. LIU, November 27, 2023



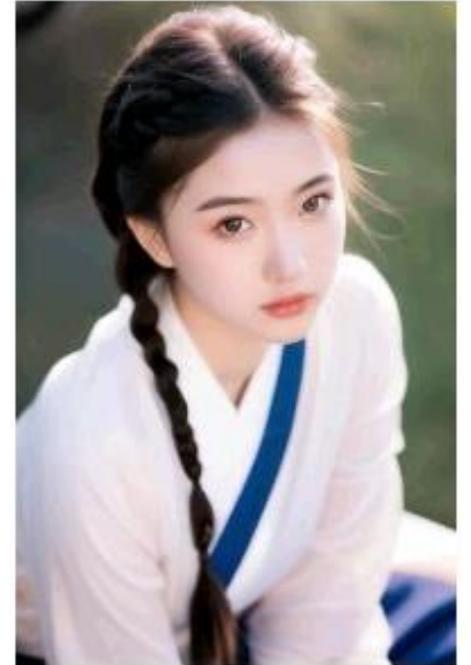
Originality

The court ruled that **the image is “fine art”** under Chinese law.

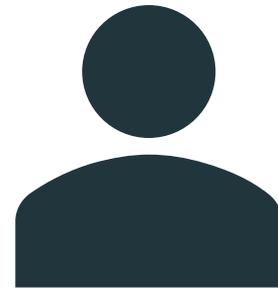
The AI did not create independently—LI made **original choices**:

- Selected AI models
- Carefully crafted prompt and negative words
- Adjusted parameters and chose the final image

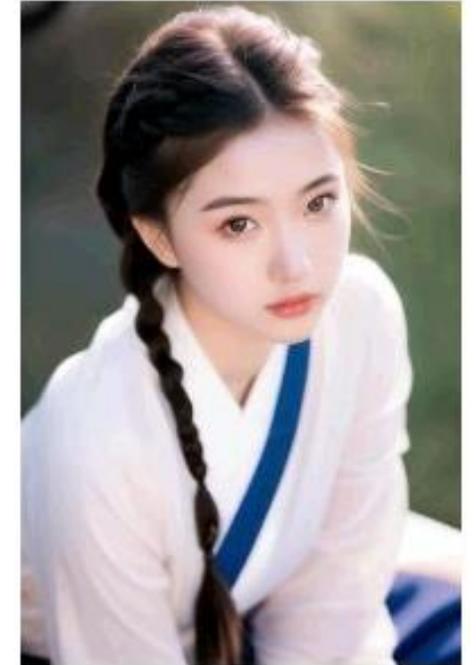
This **reflects human intellectual effort and personalised expression**.



LI v. LIU, November 27, 2023



Human



Author = **natural person** or organization.

AI cannot be the author → only a **tool**.

AI designer is not the author either → **no role in creating this image**.

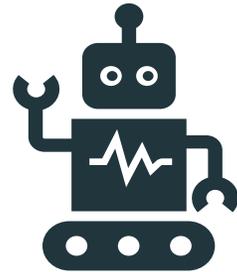
Since **LI controlled the creation process**, he is the author and owns the copyright.

LI properly disclosed AI usage by tagging “AI illustration.”

DABUS case

*German Federal Court of Justice, June 11, 2024,
file no. Z ZB 5/22*

DABUS



1

A patent application for food and beverage container was created by DABUS, an AI system.

DABUS is listed as the sole inventor.



2

The German Patent and Trademark Office rejected the application

→ Only **natural persons** can be named as inventors.

DABUS



Appeal

The German Federal Patent Court granted the patent

Condition: to list the natural person who prompted the AI system as the author



Ruling

The Federal Court of Justice upheld the decision.

DABUS



Reasoning

- AI generated inventions **can be patented** in Germany if existing criteria for protection are met.
- **Human inventor must be named** in the patent application.
- No verification on whether the named person actually made the invention

DABUS



Similar Cases

- Federal Court of Australia, Commissioner of Patents v Thaler [2022] FCAFC 62
- High Court of New Zealand, Wellington Registry, CIV-2022-485-118, [2023] NZHC 554
- South Korea, Administrative Court's Decision of June 30, 2023
- Thaler v Comptroller-General of Patents, Designs and Trademarks [2023] UKSC 49
- Japan, Tokyo District Court May 2024

To be continued

Charles-Antoine JOLY

Partner – DS AVOCATS

Intellectual Property, New Technologies and Data

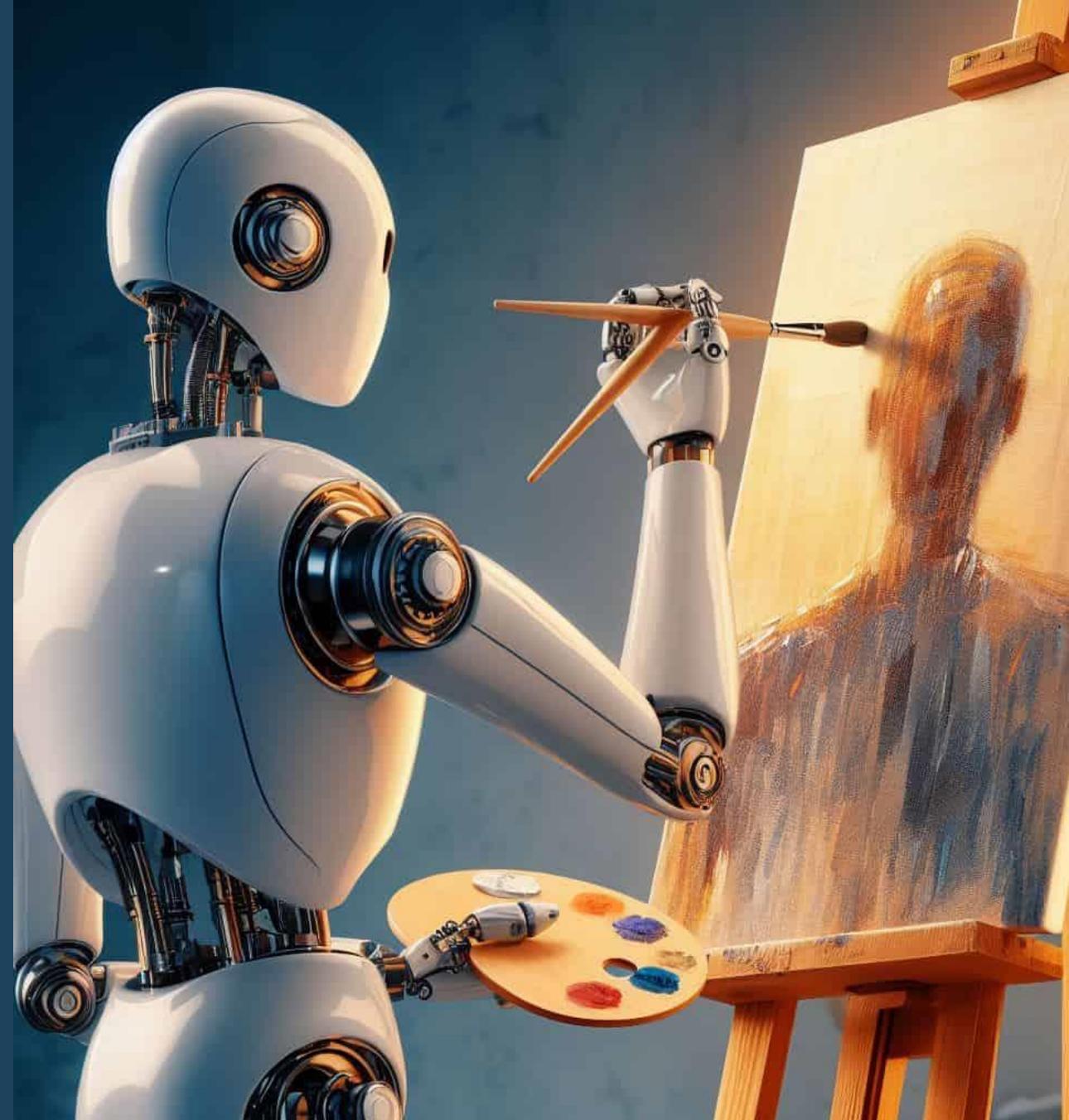
joly@dsavocats.com



**AI as a source of
infringement
The US and Chinese
perspective**

PART TWO

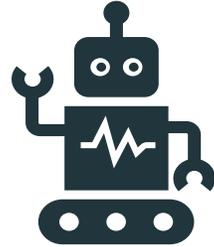
www.dsavocats.com



USA: Tremblay/Silverman v. ChatGPT

*US District Court, Case 23-cv-03223-AMO
and 23-cv-03416-AMO, Paul Tremblay et al.
V. OpenAI Inc. and al. (February 12, 2024)*

Tremblay/Silverman v. ChatGPT



1

Sarah Silverman, Paul Tremblay and others filed a lawsuit against Open AI alleging that Open AI used their works to train their AI ChatGPT without their consent.



2

Plaintiffs:

- Direct copyright infringement
- Vicarious infringement
- Violation of DMCA
- Unfair competition
- Negligence
- Unjust enrichment

Defendants:

Dismissal of all causes of action except direct copyright infringement

Tremblay/Silverman v. ChatGPT



Reasoning

Vicarious infringement:

Plaintiffs fail to explain what the outputs entail or to allege that any particular output is substantially similar – or similar at all – to their books.

Tremblay/Silverman v. ChatGPT



Reasoning

Violation of DMCA

Authors didn't prove removal of CMI: Examples of outputs

"Throughout these chapters, Tremblay masterfully maintains the suspense and psychological terror"

"Silverman uses her sharp wit to lend a comedic touch"

Plaintiffs do not point to any caselaw to suggest that failure to reveal such information has any bearing on whether the alleged removal of CMI in an internal database will knowingly enable infringement.

Tremblay/Silverman v. ChatGPT



Ruling

Dismissed claims:

- Vicarious Infringement
- Violation of DMCA
- Negligence
- Unjust Enrichment

China: LI v. LIU

*Beijing Internet Court Civil Judgment (2023),
Jing 0491 Min Chu No. 11279*

LI v. LIU, November 27, 2023



1

LI used Stable Diffusion AI software to generate a picture titled “*Spring Breeze Brings Tenderness.*”

He posted it on social media platform Little Red Book with his watermark.



2

LIU used the image without permission, removed the watermark, and published it with his own poem on another platform (Baijiahao).



3

LI sued for infringement of authorship and right to share content online, demanding a public apology and 5,000 yuan in damages.

“Spring Breeze Brings Tenderness” creation process



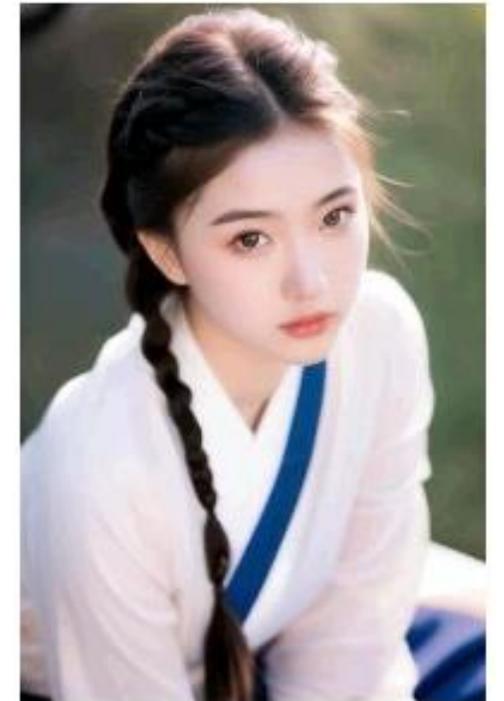
Step 6



Step 7



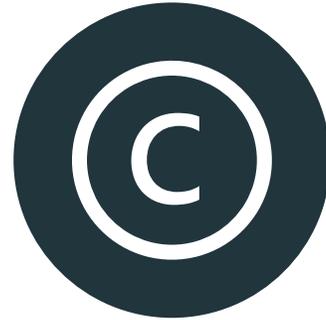
Step 8



Step 9

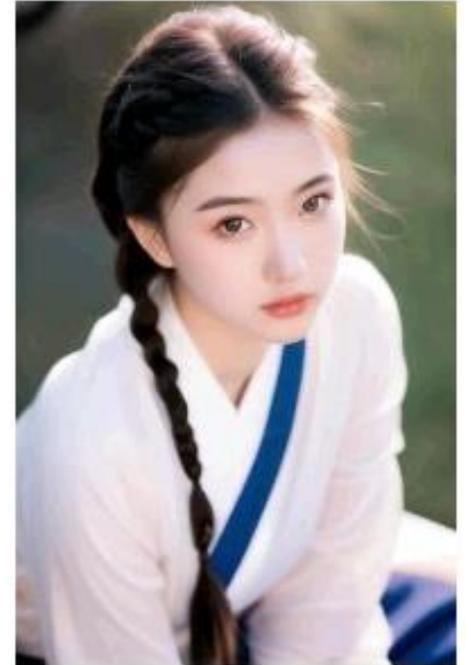
Images from LI v. LIU decision

LI v. LIU, November 27, 2023



Infringement

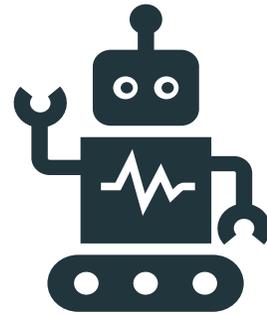
- Picture used without the author's permission.
- Watermark removed.



China: Lin Chen v. Hangzhou Gaosi Membrane Technology

*Changshu Court in Jiangsu Province,
October 18, 2024*

Lin Chen v. Hangzhou Gaosi Membrane Technology, October 18, 2024



1

Lin Chen generated an image using AI tools MidJourney and Photoshop.

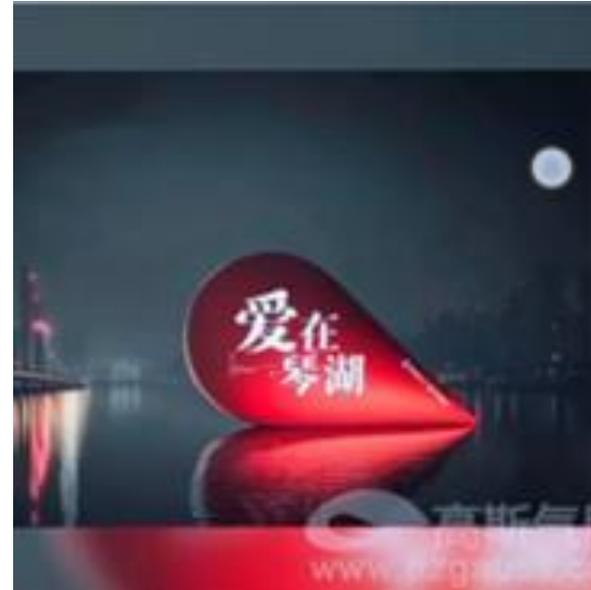
Publication on XiaoHongShu: February 14, 2023

Registration with China's Copyright Administration as a 2D artwork: April 7, 2023

Lin Chen v. Hangzhou Gaosi Membrane Technology, October 18, 2024



With Heart



**Defendant's
installation**

Images from iam-media.com

Lin Chen v. Hangzhou Gaosi Membrane Technology, October 18, 2024



2



Plaintiff:

Copyright infringement claims.

Defendant:

Gaosi, the first defendant, used the design to produce a 3D installation

QingHong, the second defendant, used the image in an advertisement

Lin Chen v. Hangzhou Gaosi Membrane Technology, October 18, 2024



Ruling

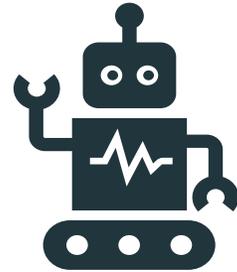


- Protection is limited to the registered 2D work.
- The defendants infringed the plaintiff's distribution and moral rights.
- The 3D installation is not an infringement.

China: Feilin v. Baidu

*Beijing Internet Court, 25 April 2019, No 239
[2019], Civil First Instance, Beijing Feilin Law
Firm v Baidu Corporation*

Feilin v. Baidu, April 25, 2019



1

The report "*Analysis Report on Judicial Big Data of the Entertainment Industry-Film Volume in Beijing*" generated by the software Wolters Kluwer China Law & Reference is published on a law firm's WeChat.

A user published the report without permission on the Baidu platform operated by the defendant.



2

Plaintiff:

Copyright infringement claims against Baidu.

Defendant:

No protection because the entire report was generated by the platform.

Feilin v. Baidu, April 25, 2019



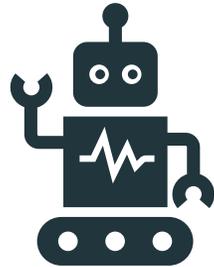
Reasoning

The analysis report, produced with the input of both the software developer and the software user, is original and can be protected even if it was generated by an IA.

China: Shenzhen Tencent v. Shanghai Yingxun

*Nanshan District People's Court,
Shenzhen, Guangdong, China [2019]:
Shenzhen Tencent v. Shanghai Yingxun,
Case No. Y0305MC No. 14010*

Tencent v. Yingxun, December 24, 2019



1

Tencent's employees use their AI Dreamwriter to publish a financial report on August 20, 2018.



2

Yingxin copied the report.

Tencent claimed copyright infringement and explained the creation process consisting of four steps.

Tencent v. Yingxun, December 24, 2019



Reasoning

Originality:

Expression, content and sentence structure = lowest degree of originality.

Author:

Not Dreamwriter: the creation process involved work from the creative team members.

→ Tencent is the copyright holder.

Tencent v. Yingxun, December 24, 2019



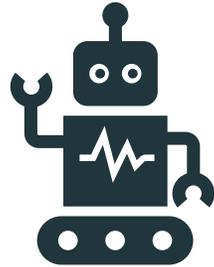
Ruling

→ Copyright infringement

China: Ultraman

- 1. Guangzhou Internet Court (2024)
Yue 0192 Min Chu 113, Ultraman*
- 2. Intermediate People's Court of Hangzhou,
Zhejiang Province*

Ultraman, February 8, 2024



1

Shanghai Character Licence
Administration holds a licence for the
Japanese series character Ultraman IP
in China from Tsuburaya Productions.



2

Plaintiff:

SCLA claimed copyright infringement
against Tab and its generative AI.

Ultraman



*Original image of Ultraman Tiga
inserted as a prompt*

Images generated by Tab

Screenshots of the Main Disputed Images Generated by Tab	
 <p>图 1 FIG. 1</p>	 <p>图 2 FIG. 2</p>
 <p>图 3 FIG. 3</p>	 <p>图 4 FIG. 4</p>
 <p>图 5 FIG. 5</p>	 <p>图 6 FIG. 6</p>

 <p>图 7 FIG. 7</p>	 <p>图 8 FIG. 8</p>
 <p>图 9 FIG. 9</p>	 <p>图 10 FIG. 10</p>

Images from Ultraman decision

Ultraman, February 8, 2024



Reasoning

Reproduction right:

- The defendant could have access to the copyright-protected Ultraman imagery
- Images generated are substantially similar to the protected work

Ultraman, February 8, 2024



Reasoning

Adaptation right:

→ The generated images partially preserved the original expression of the Ultraman model and added new features. It is an adaptation of the protected work.

→ Infringement of the right of adaptation.

Ultraman, February 8, 2024



Reasoning

Chinese AI Regulations (Interim Measures for the Management of Generative Artificial Intelligence Services):

→ Defendant is not an AI service provider, and he was required to “*take measures, such as stopping generation, stopping transmission or eliminating the illegal content, make rectification through measures such as model-based optimization training, and report findings to the relevant competent department*”

Ultraman, February 8, 2024



Appeal

Hangzhou Intermediate People's Court rejected the appeal and confirmed the judgement.

China: AI generated voice

*Beijing Internet Court (2023)
Beijing 0491 Minchu 12142 Civil Judgement
(April 23, 2024)*

Yin v. Beijing Intelligent Technology Co, April 23, 2024



1

Mr. Yin (voice actor) discovered his voice had been used in various apps without his consent.

His voice was synthesised by AI.

The AI generated voice resembled his own voice.



2

Plaintiff:

Personality rights infringement against five companies.

Yin v. Beijing Intelligent Technology Co, April 23, 2024



Reasoning

Scope of protection: whether characteristics of the voice can still be identifiable.

→ A voice synthesized using artificial intelligence can be deemed to be identifiable if it enables the general public or the public in the relevant field to associate it with the person on the basis of its timbre, intonation and pronunciation style

Yin v. Beijing Intelligent Technology Co, April 23, 2024



Ruling

- Plaintiff's voice is identifiable.
- Defendants don't have the right to authorise AI use of the voice without the plaintiff's consent.

Thank you

Charles-Antoine JOLY

Partner – DS AVOCATS

Intellectual Property, New Technologies and Data

joly@dsavocats.com

Gen AI & outputs: infringement and liability issues

Alain Strowel

Professor, UCLouvain, Magister IP/IT Alicante U, Munich IP Law Center

Attorney, Pierstone, Bruxelles

alain.strowel@uclouvain.be / alain.strowel@pierstone.com

- Co-author with Luc Desauettes and Sari Depreeuw of a study for DG Connect on AI tools and their impact on the creative sector: *Study on copyright and new technologies: copyright data management and artificial intelligence*, Publications Office of the European Union, March 2022, <https://data.europa.eu/doi/10.2759/570559> / <http://hdl.handle.net.proxy.bib.ucl.ac.be/2078/266347>

GenAI outputs: infringement and liability

- Entanglement of input & outputs issues in case law
- What kind of possible infringement (output)?
 - Outputs similar to inputs
 - **How to prove the infringement? The burden of proof (for the CJEU)**
 - Cloning of artists' styles
 - Deletion of Rights-Management Information (RMI) & trademark infringement
- Who could be liable?
 - Liability of the promptist (case of jailbreaking and RAG)?
 - Or/and of the AI developer (design of the interface and alignment)
 - + Liability of the data collector? The life cycle of AI and the training issue
- Conclusion

Copyright Lawsuits v. AI Companies

Canada

CanLII v. CasewayA
Supreme Ct. of British Columbia

Toronto Star News. v. OpenAI
Ontario Superior Ct. of J.

Gett _gg_ages US v Stabi ity AI Ltd.
High Court of Justice in London

DPG Media v. HowardsHome: TDM exception
Amsterdam Dist. Court

France

National Publishing Union v. Meta
Tribunal Judiciaire de Paris

Germany

GEMA v. OpenAI
Munich Regional Court
GEWA v. Suno
Munich Regional Court

Kneschke v. LAOIN
Hamburg Regional Court

South Korea

Korean Broad. Assn. v. Naver
Seoul Central Dist. Court

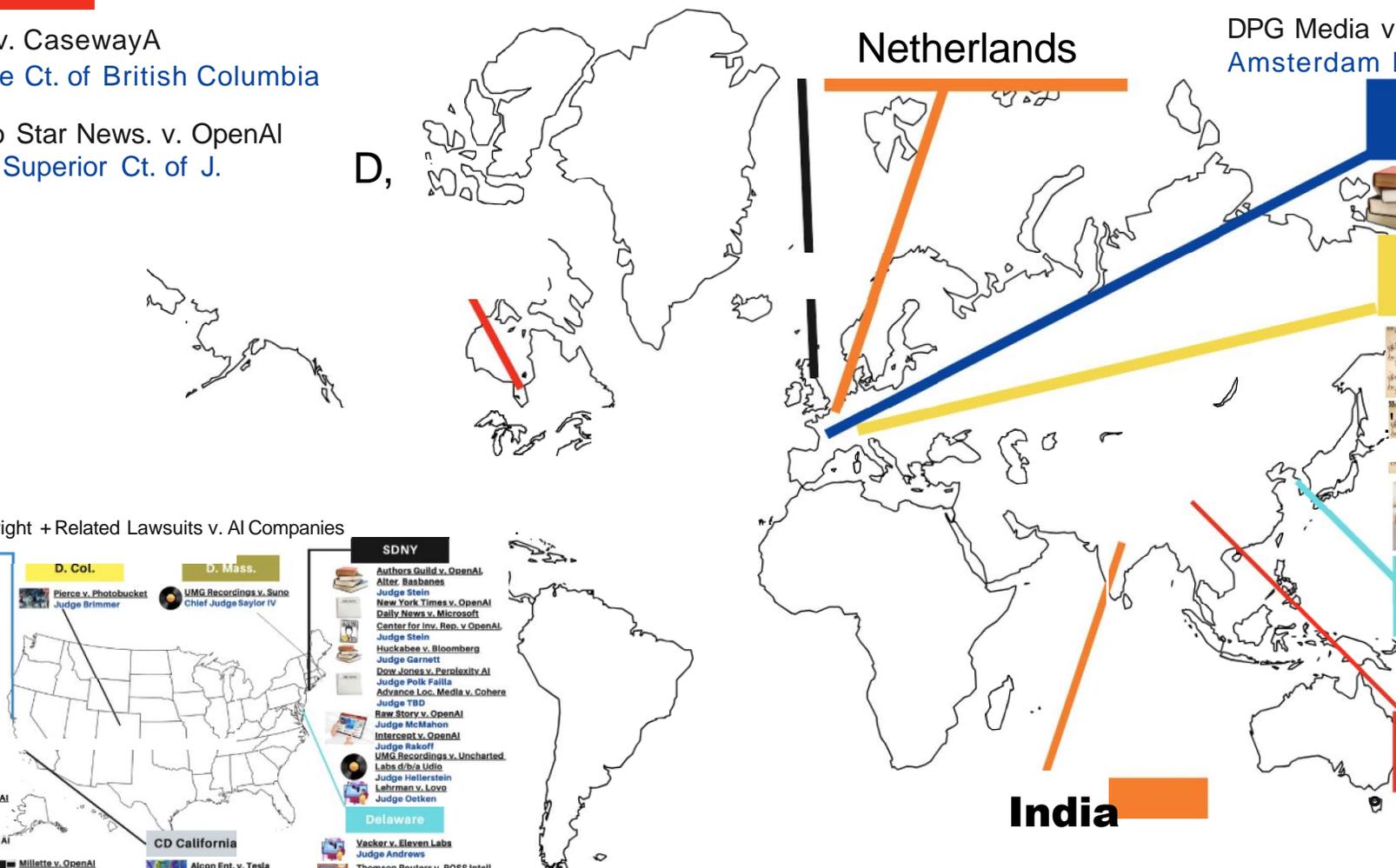
China

Shanghai Character License Administrative Co. v. TAB
Guangzhou Internet Court

Netherlands

India

Asian News Int'l. v. OpenAI
New Delhi Court



Copyright + Related Lawsuits v. AI Companies

ND California Kadrey v. Meta; Chabon v. Meta Huckabee v. Meta; Earnsworth Judge Chhabria In re OpenAI ChatGPT Litigation (Tremblay, Silverman, Chabon) Judge Araceli Martinez-Olguin Nazemian v. NVIDIA Corp. Dubau v. NVIDIA Corp. (related) Judge Tiger In re Mosaic LLM Litigation (O'Nan; Makka) Judge Breyer Sarav v. Anthropic Concord Music v. Anthropic Judge Leo Doe 1 v. Github; Microsoft; OpenAI Judge Tiger Sarah Andersen v. Stability AI Deviant Art, Midjourney, Runway AI Judge Orrick In re Google Gen. AI Ltn. (Zhang; Leovy) Judge Lee Brave Software v. News Corp. Judge Breyer	D. Col. Pierce v. Photobucket Judge Brimmer	D. Mass. UMG Recordings v. Suno Chief Judge Saylor IV	SDNY Authors Guild v. OpenAI Alter; Basbanes Judge Stein New York Times v. OpenAI Daily News v. Microsoft Center for Inv. Rep. v. OpenAI Judge Stein Huckabee v. Bloomberg Judge Garnett Dow Jones v. Perplexity AI Judge Polk Falls Advance Loc. Media v. Cohere Judge TBD Raw Story v. OpenAI Judge McMahon Intercept v. OpenAI Judge Rakoff UMG Recordings v. Uncharted Labs d/b/a Ludio Judge Hellerstein Lehman v. Levo Judge Oetken	Delaware Vacker v. Eleven Labs Judge Andrews Thomson Reuters v. ROSS Intell. Judge Bibas Getty Images v. Stability AI Judge Hall
CD California Millette v. OpenAI Millette v. Google Millette v. NVIDIA Judge Donato	Alcon Ent. v. Tesla Judge Oliver			



Copyright > Related Lawsuits v. AI Companies



Kadrey Meet Chabon v. Meta
 eyue v Me Farnsworth
 Judge Chhabria



In re OpenAI ChatGPT Litigation
 (Trembfay, vèSlrma_,n Chabo}n
 Judge Araceli Martinez-Olgufn



Nazemian v. NVIDIA Corp
 Dubus v. NVIDIA Corp (related)
 Judge Tigar



In re Mosaic LLM Litigation
 {U'Nan; Makkai)
 Judge Breyer



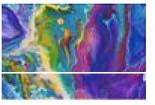
Bartz v. Anthropic
 Judge Alsup



Concord Music. v. Anthropic
 Judge Lee



Doe 1 v. Github, Microsoft, OpenAI
 Judge Tigar



Sarah Andersen v. **StabilityAI** ..
 Deviant Art, Midjourney, Runway Ai
 Judge Orrick



In re Google Gen. AT Ltgn
 (Zhang, Leovy)
 Judge Lee



Brave Software v. News Corp
 Judge Breyer

D. Col.

Pierce v. Photobucket
 Judge Brimmer

D. MASS.

UMG Recordings v. Suno
 Chief Judge Saylor IV



Authors Guild v. Opne, AI
 AI r,e Basbanes

Judge Stein
 New York Times v. OpenAI

Dashy News v. Microsoft

Center for Inv. Rep. v OpneAI
 Judge Stein

Huckabee v. Bloomberg

Judge Garnett

Dow Jones v. Perplexity A

Judge Potk Faila

Advance Loc. Media v. Cohere

Judge TBD

Raw Story v. OpenAI

Judge McMahon

Intercept v. OpenAI

Judge Rakoff

UMG Recordings v. Uncharted

Labs d/b/a Udio

Judge Hellerstein

Lehrman v. Lovo

Judge Oetken



"



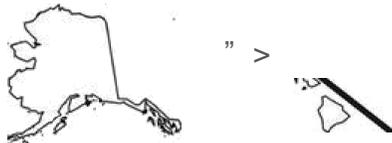
Vacker v. Eleven Labs
 Judge Andrews



Thomson Reuters v. ROSS Intels.
 Judge Bibas



Getty_fllt_ages v. **Stability AI**
 Judge Half



CD California



Alcon Ent. v. Tesla
 Judge Oliver



Millette v. OpenAI
 Millette v. Google
 Mittette v. NVIDIA
 Judge Donato

Claims in the 40+ US cases

- US Class actions filed by authors since 2023:
 - Jan.: against ***GitHub and Microsoft***: alleged violation of **developers' copyright + Digital Millenium Copyright Act or DMCA violations** (partly dismissed in May 2023)
 - Jan.: ***Sarah Andersen etc v. Stability AI, DeviantArt and Midjourney***: artists claim these companies have without authorization used > 5 billion works (LAION = dataset constituted by a German non-profit), including their own, and have stored them to produce **derivative works (copyright)** (partly confirmed/dismissed in July 2024)
 - June (***Paul Tremblay v. OpenAI***) + July (***Sarah Silverman etc. v. OpenAI & Meta*** (LLaMa)): remix of thousands of works without consent, compensation, or credit + Sept (***Author's Guild, George Martin, John Grisham etc. v. OpenAI***): for **direct** and **vicarious © violations**
- Actions by companies:
 - ***Getty Images v. Stability AI*** (29/3/23 in US+UK): the processing of images (+metadata) to train = © infringement + generated images contain a **modified version of the watermark**, creating confusion as to the source of the images (**copyright + trademark**)
 - ***New York Times v. OpenAI/Microsoft*** (27/12/23) → damages+destruction of dataset/GPTs
 - Other Newspapers cases: ***Daily News, Center for Investigative Reporting***

No US decision on the merits, but some trends

- More peripheral claims tend to be dismissed:
 - Unfair competition (California Act)
 - Negligence (no respect of a general duty to care)
 - Unjust enrichment → idea of (unjust) taking
 - Some DMCA objections
- Other claims remain for consideration for ex. in the 3 newspapers' cases (*NYT, Daily News, Center for Investigative Reporting or CIR*):
 - Direct copyright claim
 - Copyright contributory infringement
 - Trademark claims (of dilution)
 - A DMCA claim against OpenAI in *Daily News* and *CIR*

US: copyright claims turn around fair use (s. 107 US Copyright Act)

The conditions of *fair use* are assessed through a four-step test, which considers the following factors:

1. The purpose and character of the use,
Including whether it is commercial or not, educational or critical.
2. The nature of the protected work
For example, whether it is more creative or more informative.
3. The amount and quality of use of copyrighted material.
A small portion of the work, used in an educational or critical context, is more likely to be considered fair use.
4. The potential effect of the use on the market or value of the copyrighted work.

A precedent in the US: the *Google Books* case refers to the outputs: displayed « short snippets »

"Google's unauthorized digitization of copyrighted works, creation of a search feature, and display of short snippets of those works are non-infringing fair uses. The purpose of the copying is highly transformative, the public display of the text is limited, and the disclosures do not provide a meaningful commercial substitute for the protected aspects of the originals. Google's commercial nature and profit motive do not justify denying fair use."

(Authors Guild v. Google, 804 F3d 202, 229 (2d Cir 2015))

A relevant precedent?

Possible major difference: the excerpts of books did not create a significant market substitute for the books.

Does Gen AI replace the source? Does it generate competing content?

Entanglement shown in *Thomson v. Ross Intelligence* re a legal AI tool (summary judgment, 11 Febr. 2025)

- Seems to focus on the input/training phase
 - Thus on the copying of original elements
 - Actual copying: Ross had access to Westlaw and used Thomson's work to make the Bulk Memos
 - Substantial similarity
- But in the discussion of the defense of fair use:
 - burden of proof for Ross/defendant
 - 4 factors analysed, here 3 (for Ross) and 4 (for Thomson as well as overall balancing):
 - How much was used/how substantial: what matters = « not the amount and substantiality of the portion used in making the copy, but rather the amount and substantiality of what is thereby made accessible to a public » → based on the similarity in outputs: whether « competing substitute » or not
 - Effect on the market/value of work (most important factor assessed with regard to outputs)
 - But including on the potential derivative markets = « data to train legal AIs » = input consideration

Output-input entanglement shown in *Thomson Reuters v. Ross Intelligence* (11 Febr. 2025)

Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345 (1991).

Question

Does originality for copyright purposes mean that the work was independently created and has some minimal degree of creativity?

West Headnote

Originality, for copyright purposes, means that the work was independently created and has some minimal degree of creativity.

Case Opinion

Original, as the term is used in copyright, means only that the work was independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity.

GenAI outputs: infringement and liability

- Entanglement of input & outputs issues in case law
- What kind of possible infringement (output)?
 - 1. Outputs similar to inputs
 - **How to prove the infringement? The burden of proof (for the CJEU)**
 - 2. Cloning of artists' styles
 - 3. Deletion of Rights-Management Information (RMI) & trademark infringement
- Who could be liable?
 - Liability of the promptist (case of jailbreaking and RAG)?
 - Or/and of the AI developer (design of the interface and alignment)
 - + Liability of the data collector? The life cycle of AI and the training issue
- Conclusion

What kind of infringement?

1. AI outputs look similar to inputs



Image Cornell U (from weforum.org)

ARTIFICIAL INTELLIGENCE

AI models spit out photos of real people and copyrighted images

The finding could strengthen artists' claims that AI companies are infringing their rights.

By **Melissa Heikkilä**

February 3, 2023



Extracts from the New York Times complaint against OpenAI & Microsoft

Output from GPT-4:

exempted it from regulations, subsidized its operations and promoted its practices, records and interviews showed.

Their actions turned one of the best-known symbols of New York — its yellow cabs — into a financial trap for thousands of immigrant drivers. More than 950 have filed for bankruptcy, according to a Times analysis of court records, and many more struggle to stay afloat.

“Nobody wanted to upset the industry,” said David Klahr, who from 2007 to 2016 held several management posts at the Taxi and Limousine Commission, the city agency that oversees medallions. “Nobody wanted to kill the golden goose.”

New York City in particular failed the taxi industry, The Times found. Two former mayors, Rudolph W. Giuliani and Michael R. Bloomberg, placed political allies inside the Taxi and Limousine Commission and directed it to sell medallions to help them balance budgets and fund key initiatives.

During that period, much like in the mortgage lending crisis, a group of industry leaders enriched themselves by artificially inflating medallion prices. They encouraged medallion buyers to borrow as much as possible and ensnared them in interest-only loans and other one-sided deals that often required borrowers to pay hefty fees, forfeit their legal rights and give up most of their monthly incomes.

When the market collapsed, the government largely abandoned the drivers who bore the brunt of the crisis. Officials did not bail out borrowers or persuade banks to soften loan

Actual text from NYTimes:

exempted it from regulations, subsidized its operations and promoted its practices, records and interviews showed.

Their actions turned one of the best-known symbols of New York — its signature yellow cabs — into a financial trap for thousands of immigrant drivers. More than 950 have filed for bankruptcy, according to a Times analysis of court records, and many more struggle to stay afloat.

“Nobody wanted to upset the industry,” said David Klahr, who from 2007 to 2016 held several management posts at the Taxi and Limousine Commission, the city agency that oversees cabs. “Nobody wanted to kill the golden goose.”

New York City in particular failed the taxi industry, The Times found. Two former mayors, Rudolph W. Giuliani and Michael R. Bloomberg, placed political allies inside the Taxi and Limousine Commission and directed it to sell medallions to help them balance budgets and fund priorities. Mayor Bill de Blasio continued the policies.

Under Mr. Bloomberg and Mr. de Blasio, the city made more than \$855 million by selling taxi medallions and collecting taxes on private sales, according to the city.

But during that period, much like in the mortgage lending crisis, a group of industry leaders enriched themselves by artificially inflating medallion prices. They encouraged medallion buyers to borrow as much as possible and ensnared them in interest-only loans and other one-sided deals that often required them to pay hefty fees, forfeit their legal rights and give up most of their monthly incomes.

Burden of proof when output similarities

- For the plaintiff to prove the « free and creative choices » (CJEU, *Painer* etc.)
 - + to identify the relevant elements (CJEU, *Karen Millen*, pt 47 for individual character of an unregistered design)
 - Novelty (no precedent) points towards originality
- When similarities, need to prove copying (>< patent) = not an independent creation: harmonized by CJEU (for unregistered design, *Gautzch*, pt 44 ff)
 - If impossible or excessively difficult → possible mitigation/lightening of burden
- Presumption of copying if (substantial) similarities or if possibility of accessing the protected work
 - Then for the defendant to prove independent creation
- « Fair balance » between fundamental rights for protection & infringement

2. New questions on the protection of style



(Research by the University of Tübingen)

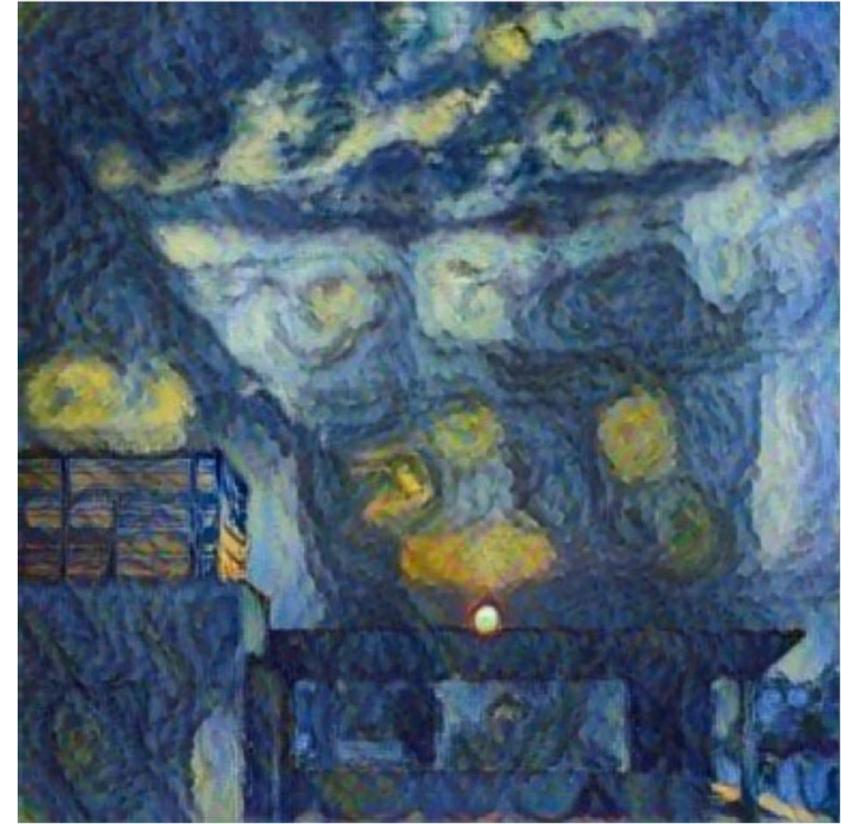
Refusal of US © Office to register (Review Board)

- Case *Suryast* (decision 9 Dec. 2023): painting in the Van Gogh style by the AI RAGHAV
- « *Mr. Sahni generated the Work by taking an original photograph that he authored, inputting that photograph into RAGHAV, then inputting a copy of Vincent van Gogh's The Starry Night into RAGHAV as the "style" input to be applied to the photograph, and choosing "a variable value determining the amount of style transfer."* Mr. Sahni further explained that he named RAGHAV as a co-author because its "contribution is distinct, disparate and independent" from his contribution to the Work »
- Not enough human content in the derived work (comp. digital assistant for adjusting a photo)



+

Image of a style (The Starry Night)





25/3/2025: new OpenAI image generator replacing DALL-E (integrated in GPT-4o)

A Ghibli effect? Flood of images in the style of the Japanese studio Ghibli



OpenAI view, reactions and old views

- **Response of OpenAI:** « “We continue to prevent generations in the **style of individual living artists, but we do permit broader studio styles** — which people have used to generate and share some truly delightful and inspired original fan creations.”
- **Replies:** story board artist (about LensaAI): “But I would argue that for us, **our style is actually our identity**. It’s is what sets us apart from each other. It’s what makes us marketable to clients.”
- **Old views on style:**
 - « Le style est l’homme même » (Buffon)
 - « Le style, c’est l’homme: quand le style est obscur, il faut déjà s’inquiéter » (A. Comte-Sponville)
 - « Quand on voit le style naturel, on est tout étonné et ravi, car on s’attendait de voir un auteur, et on trouve un homme » (Blaise Pascal)
 - « Le style est une manière de dire qui vous êtes sans parler » (Rachel Zoé, styliste)



Greg Rutkowski, a Polish artist who specializes in fantastical scenes. His name has become shorthand for users who wanted to generate fanciful images. Dorota Rutkowska

Stable Diffusion 1 Demo

Stable Diffusion is a state of the art text-to-image model that generates images from text. For faster generation and API access you can try [DreamStudio Beta](#)

dragon by greg rutkowski

Generate image



MIT Technology Review Subscribe

"Dragon Cave"
GREG RUTKOWSKI

ARTIFICIAL INTELLIGENCE

This artist is dominating AI-generated art. And he's not happy about it.

Greg Rutkowski is a more popular prompt than Picasso.

by **Melissa Heikkilä**

September 16, 2022

About Greg Rutkowski

Independent graphic artist

Possible to generate outputs, through prompts referring to his style?

I ask for an image in the style of Greg Rutkowski

pourquoi n'est-ce pas possible?

La création d'images directement inspirées par le style spécifique d'artistes contemporains ou bien protégés par des droits d'auteur pose des problèmes de respect des droits de propriété intellectuelle. Cela inclut des artistes vivants ou dont l'œuvre est encore sous droit d'auteur. Je suis là pour aider à créer quelque chose d'unique pour vous en m'inspirant de styles généraux ou de concepts, plutôt que de reproduire exactement le style d'un artiste spécifique. Si vous avez d'autres idées ou si vous souhaitez explorer un style ou un thème général, je serais ravi de vous aider à conceptualiser et créer quelque chose de nouveau !

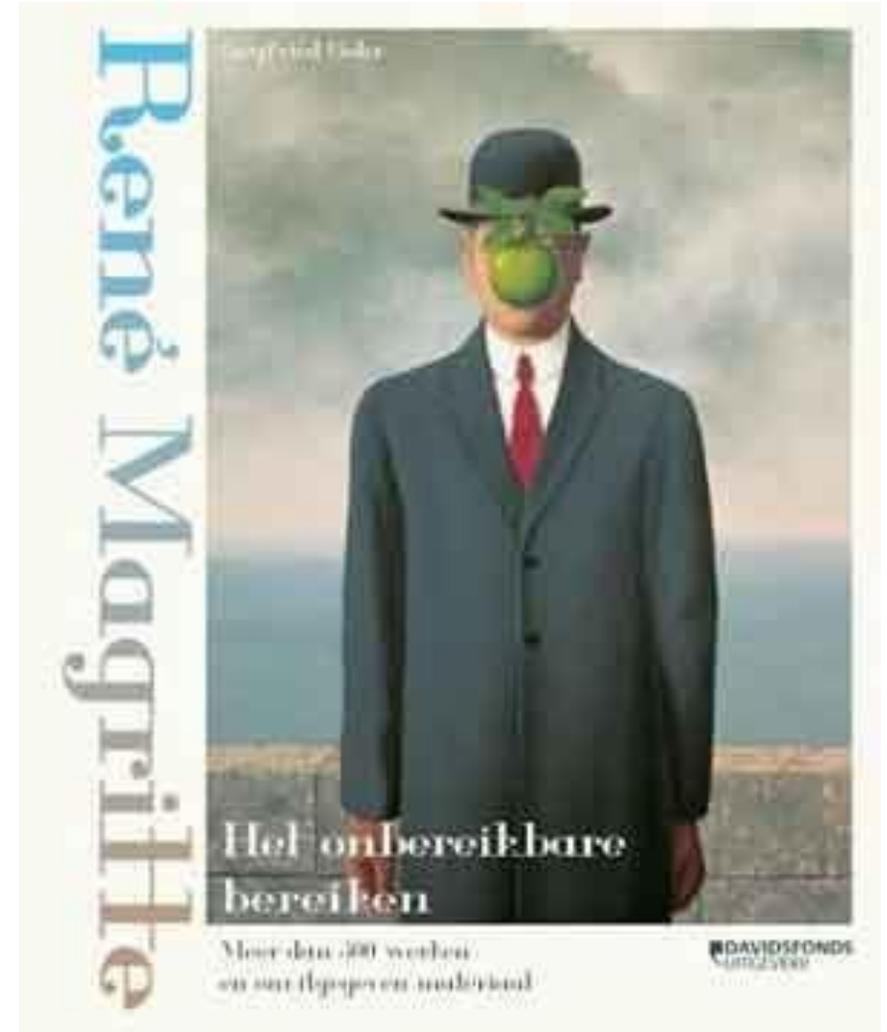


Appropriation of style

- Standard view: a style is not protected
 - Why? Probably because it is **difficult to define and grasp**.
 - Also because it is **too abstract**: it is the manner, way, or method in which some artisanal or craft practice is performed
 - Because style is **equated with a genre** (collage or impressionist style, etc.)
 - And because **it is not fixed** and not clearly identifiable. But we recognize a style, it is an ineffable expression of personality.
- Other views: distinction of group (or communal) style: not protected, and individual (or personal) style: can be protected
 - When the public is led to believe that it comes from an individual artist (false indication → free-riding)
- Wisdom (but 'duplicity'?) of GPT-4o:
 - Distinction between specific and general style: the style of an artist is specific enough!
 - Alignment: no output if the prompt asks for a specific style of a living author or of an author's whose works are still protected

Créer une image peux-tu créer une image dans le style de Magritte

Image créée

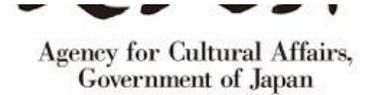


AI tools (cl)aim to capture personal styles (and sometimes offer a remuneration)

- **Grammarly:** a cloud-based typing assistant, has launched a “personalized voice detection and application” which automatically detects a person’s unique writing style and creates a “voice profile” that can rewrite any text in the person’s style.
 - Creepy?!
- **Tess AI:** remunerates artists for the copying of their style

Tess is the world's first properly-licensed **AI image generator**. Our mission is to empower creative people to leverage AI ethically. To this end, we've built **a platform that allows creators to generate images in a consistent visual style**, and for the **artists behind the styles to be fairly compensated for their work.**

PROTECTION OF STYLE IN JAPON ?



“General Understanding on AI and Copyright in Japan” -Overview-

(published by the Legal Subcommittee under the Copyright Subdivision of the Cultural Council)

Japan Copyright Office (JCO)
(Copyright Division, Agency for Cultural Affairs, Japan)

May 2024

Imitation of a "Creator's Style"

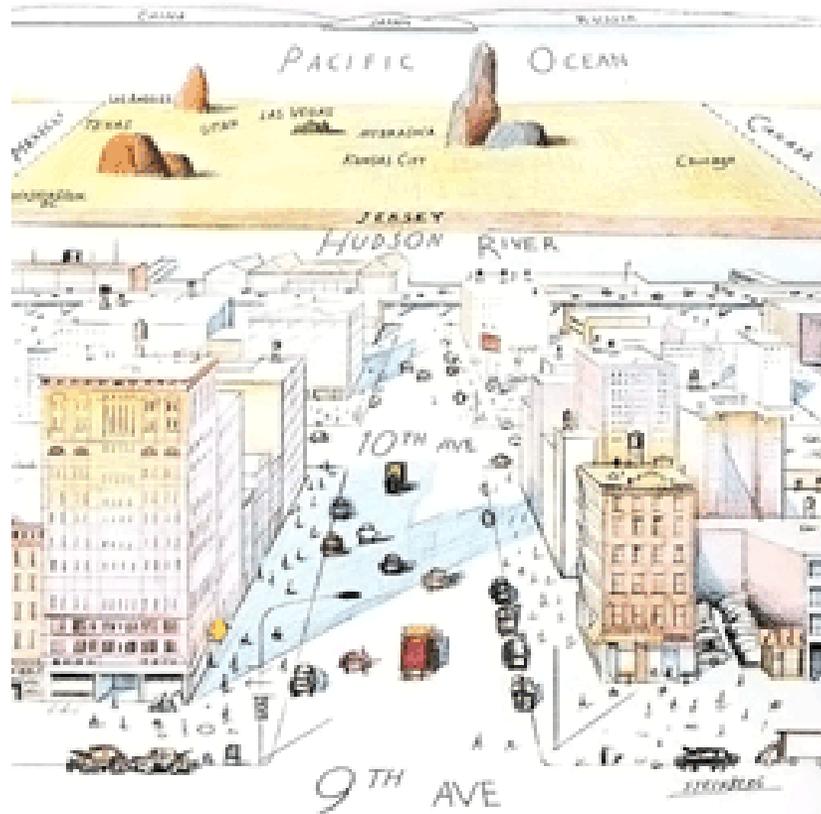
- ❑ Copyright protection applies to the "creative expressions" of an idea and not to the "idea" itself. Therefore, an AI-generated material which applies the "creator's style" of a preexisting copyrighted work does not infringe copyright if the style merely encompasses an idea.
- ❑ Conversely, Article 30-4 of the Act may not apply to AI training which "picks off" a specific creator's style in the following situations:

- The distinction between "ideas" and "creative expressions" has been determined on a case-by-case basis, depending on the specific circumstances.
- A small group of works consisting only of specific creator's copyrighted works may not only have a common "creator's style" (i.e., idea), but may also include common "creative expressions."
- If such a group of works is intentionally reproduced in order to carry out additional training for the purpose of generating materials that contains all or parts of such common creative expressions, Article 30-4 of the Act does not apply to such reproduction because the "purpose of enjoyment" coexists.

Few precedents on the protection of style

- *Steinberg v. Columbia Pictures*, 1987, US District Court, S.D. New York

Mar. 29, 1976 Price 75 cents
THE NEW YORKER



Steinberg v. Columbia Pictures, 1987, US District Court, S.D. New York

" **The "Moscow" illustration** [...] : is **executed** in a **blend of styles**: the three characters, whose likenesses were **copied from a photograph**, have realistic faces and somewhat sketchy clothing, and the **city blocks are drawn in a fairly detailed but sketchy style.** "

" one can see the **striking stylistic relationship between the posters**, and since **style is one ingredient of "expression,"** this relationship is significant. Defendants' illustration was executed in the sketchy, whimsical style that has become one of Steinberg's hallmarks "

To go further:
comparison
US, EU, Japan:
Report GPAI /
OCDE

Co-generation of data

Copyright and Data Protection Rights in
Co-Generated Input and Output of
Generative AI

November 2024



This report was planned prior to the integration of the Global Partnership on Artificial Intelligence (GPAI) and the Organisation for Economic Co-operation and Development (OECD) mid-2024. Consequently, the report was not subject to approval by GPAI and OECD members and should not be considered to reflect their positions.

3. Removing right-management information (RMI)



Figure 2: Exhibit from *Getty Images Inc. v. Stability AI, Inc.* District Court of Delaware, 2023.

Gemini 2.0 flash to edit pictures:
with a simple prompt, you erase
the watermark (Techcrunch,
16/3/2025)

>< GPT-4o et Claude 3.7

People are using Google's new
AI model to remove
watermarks from images

Kyle Wiggers — 1:49 PM PDT · March 17, 2025



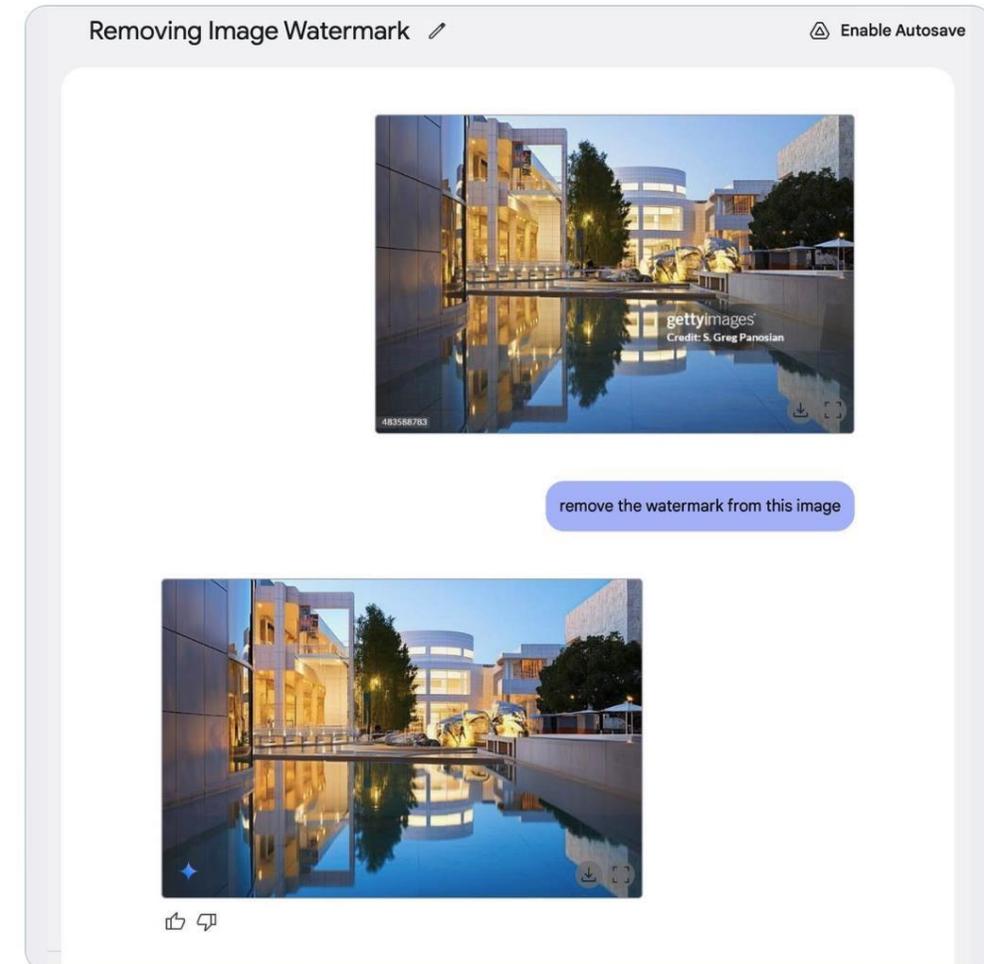
Tanay Jaipuria

@tanayj · Follow



Gemini 2.0 Flash, available in Google's AI studio, is amazing at editing images with simple text prompts.

It also can remove watermarks from images (and puts its own subtle watermark in instead 🌀)



8:58 PM · Mar 16, 2025



Protection of Rights-management information or RMI (and TPM)

- Art. 7 of 2001/29 InfoSoc dir. mandates « adequate legal protection against any person knowingly performing without authority any of the following acts:
 - (a) the removal or alteration of any electronic rights-management information;
 - (b) the distribution, importation for distribution, broadcasting, communication or making available to the public of works (...) **from which electronic rights-management information has been removed or altered without authority »**
- Article 6 mandates « adequate legal protection against the circumvention of any effective technological measures (TPM), which the person concerned carries out in the knowledge, or with reasonable grounds to know, that he or she is pursuing that objective »
- TPM = « any technology, device or component that, in the normal course of its operation, is designed to prevent or restrict acts, in respect of works or other subject-matter, which are not authorised »

GenAI outputs: infringement and liability

- Entanglement of input & outputs issues in case law
- What kind of possible infringement (output)?
 - 1. Outputs similar to inputs
 - **How to prove the infringement? The burden of proof (for the CJEU)**
 - 2. Cloning of artists' styles
 - 3. Deletion of Rights-Management Information (RMI) & trademark infringement
- Who could be liable?
 - Liability of the promptist (case of jailbreaking and RAG)?
 - Or/and of the AI developer (design of the interface and alignment)
 - + Liability of the data collector? The life cycle of AI and the training issue
- Conclusion

Who is liable?

Possible sole responsibility of users

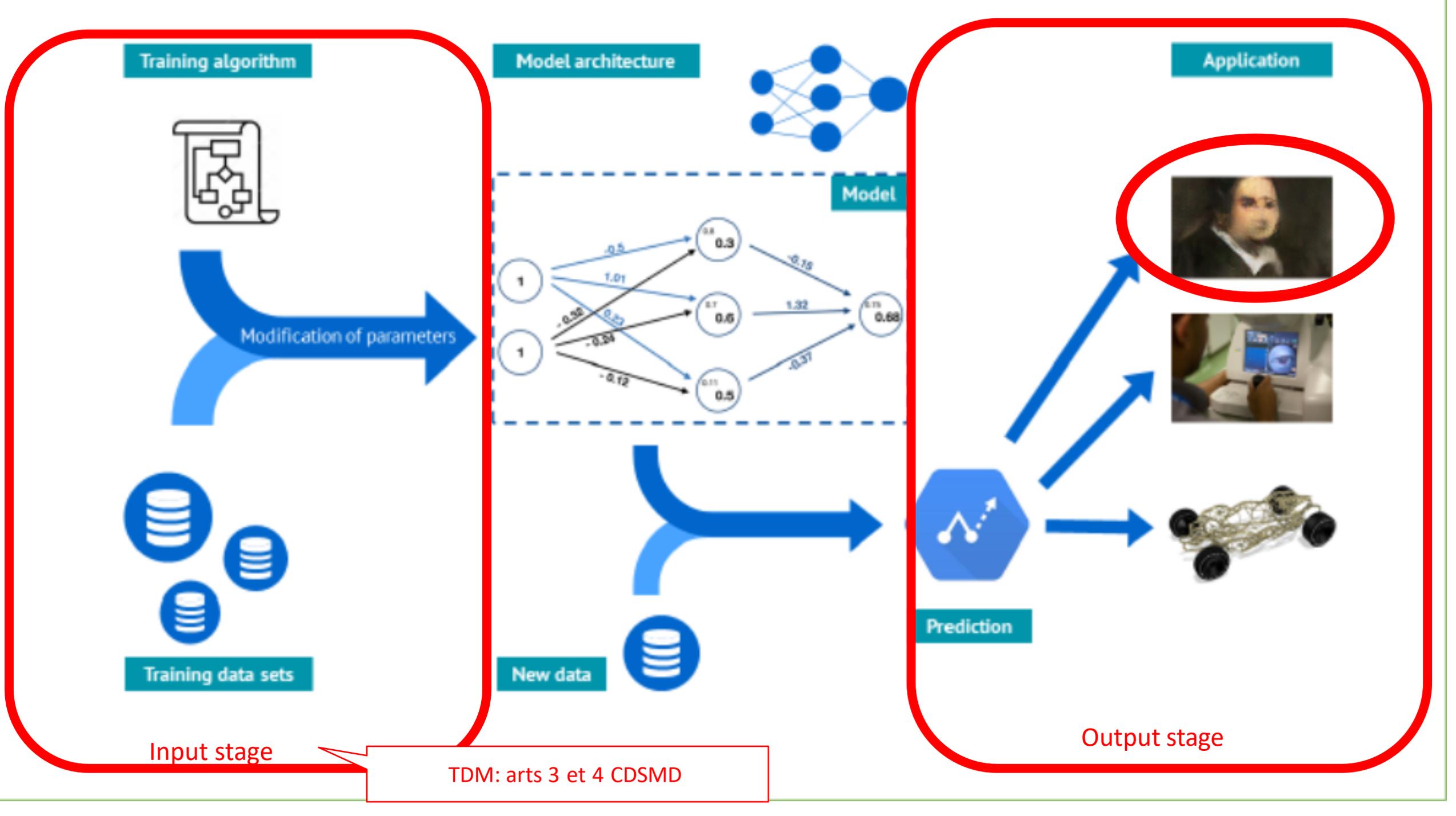
- **Jailbreaking** « refers to the process of manipulating a GenAI model to bypass its built-in safety measures and produce unintended outputs through carefully crafted prompts. This vulnerability can arise from either architectural limitations or training data biases »
- **RAG** (Retrieval Augmented Generation): aims to optimize the outputs by connecting an AI model with an external database
 - What right does the user have on what s/he inputs?

Research [AI and machine learning](#) [AI threats](#)

7 min read

AI jailbreaks: What they are and how they can be mitigated

By [Microsoft Threat Intelligence](#)



Life cycle: training data 'locked' into model weights
making it possible for outputs to imitate the inputs

- Probabilities in the transformer model (like GPTtransformer), not natural language in human-readable form
- For Gen models based on the transformer architecture, memorization is undesired and a bug
 - But AI models are prone to memorize fragments of the training data when they appear often (ex the press articles shared multiple times by the users of Reddit), and those fragments might contain verbatim reproductions of works (without attribution)
 - Deduplication is not sufficient
- The alignment phase is also important: possibility to prohibit or not certain output (not doing it might be considered as inducing copyright infringement - see inducement theory for P2P systems, cf. *Grokster* US SC but *Rapidshare* case law in SW/CH: not an accomplice)
 - Not (only) the liability of the user (but joint responsibility)

Dernier mot pour ChatGPT sur le Text and Data Mining for research (art. 3 Dir. 2019/790): étonnant!

En ce qui concerne l'utilisation de données protégées par le droit d'auteur pour entraîner des systèmes d'IA, l'Union européenne a récemment adopté la Directive sur le droit d'auteur dans le marché unique numérique, qui contient une exception pour l'utilisation de certaines œuvres protégées par le droit d'auteur à des fins de text et data mining, qui est utilisé pour entraîner des systèmes d'IA. **Cependant, cette exception est limitée aux fins de recherche scientifique et est soumise à certaines conditions, telles que l'obligation de fournir une rémunération équitable aux titulaires de droits concernés.**

Pas sûr que OpenAI soit du même avis !!

peux-tu comparer l'application des exceptions au droit d'auteur en droit américain et en droit européen au cas où un outil d'intelligence artificielle utilise des données de départ pour entraîner le système, parmi lesquelles on trouve des données comme des photos ou du texte qui sont protégés par un droit d'auteur?

Aux États-Unis, le droit d'auteur contient une exception de "fair use" qui permet l'utilisation d'œuvres protégées par le droit d'auteur à des fins telles que la critique, les commentaires, l'enseignement, les rapports d'actualités ou la recherche scientifique. Cette exception est appliquée en cas d'utilisation d'œuvres protégées par le droit d'auteur dans un contexte d'IA pour entraîner un système d'IA. La Cour suprême américaine a également récemment confirmé dans l'affaire Google c. Oracle que l'utilisation de certaines parties d'une œuvre protégée par le droit d'auteur à des fins d'interopérabilité avec d'autres logiciels était un "fair use".

En Europe, le droit d'auteur contient également certaines exceptions telles que la citation, la parodie, la caricature et l'enseignement, mais ces exceptions sont plus limitées que la notion de "fair use" aux États-Unis. En ce qui concerne l'utilisation de données protégées par le droit d'auteur pour entraîner des systèmes d'IA, l'Union européenne a récemment adopté la Directive sur le droit d'auteur dans le marché unique numérique, qui contient une exception pour l'utilisation de certaines œuvres protégées par le droit d'auteur à des fins de text et data mining, qui est utilisé pour entraîner des systèmes d'IA. Cependant, cette exception est limitée aux fins de recherche scientifique et est soumise à certaines conditions, telles que l'obligation de fournir une rémunération équitable aux titulaires de droits concernés.

Thanks for your attention

More on:

AI and the
law
AI for lawyers

Alain Strowel

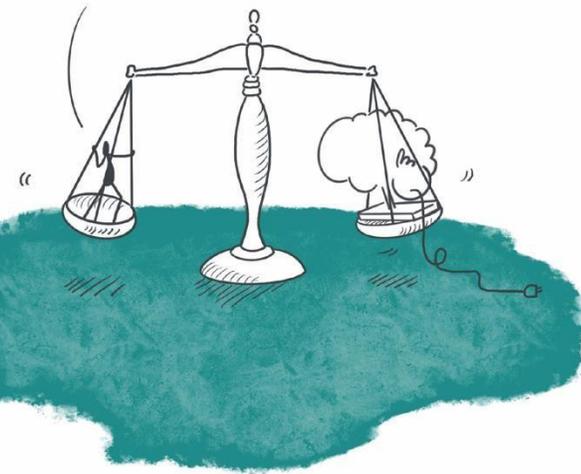
alain.strowel@uclouvain.be

Alain.strowel@pierstone.com

L'INTELLIGENCE ARTIFICIELLE POUR LES JURISTES

ALAIN STROWEL ET FRANÇOIS WÉRY

TU SAIS C'EST
JUSTE UNE QUESTION
D'ÉQUILIBRE



 LARCIER
INTERSENTIA

Gen AI & outputs: infringement and liability issues

Alain Strowel

Professor, UCLouvain, Magister IP/IT Alicante U, Munich IP Law Center

Attorney, Pierstone, Bruxelles

alain.strowel@uclouvain.be / alain.strowel@pierstone.com

- Co-author with Luc Desaunettes and Sari Depreeuw of a study for DG Connect on AI tools and their impact on the creative sector: *Study on copyright and new technologies: copyright data management and artificial intelligence*, Publications Office of the European Union, March 2022, <https://data.europa.eu/doi/10.2759/570559/> / <http://hdl.handle.net.proxy.bib.ucl.ac.be/2078/266347>

GenAI outputs: infringement and liability

- Entanglement of input & outputs issues in case law
- What kind of possible infringement (output)?
 - Outputs similar to inputs
 - **How to prove the infringement? The burden of proof (for the CJEU)**
 - Cloning of artists' styles
 - Deletion of Rights-Management Information (RMI) & trademark infringement
- Who could be liable?
 - Liability of the promptist (case of jailbreaking and RAG)?
 - Or/and of the AI developer (design of the interface and alignment)
 - + Liability of the data collector? The life cycle of AI and the training issue
- Conclusion

Copyright Lawsuits v. AI Companies

Canada



CanLII v. Caseway AI
Supreme Ct. of British Columbia



Toronto Star News. v. OpenAI
Ontario Superior Ct. of J.

UK



Getty Images (US) v. Stability AI Ltd.
High Court of Justice in London

Netherlands



DPG Media v. HowardsHome: TDM exception
Amsterdam Dist. Court

France



National Publishing Union v. Meta
Tribunal Judiciaire de Paris

Germany



GEMA v. OpenAI
Munich Regional Court
GEMA v. Suno
Munich Regional Court



Kneschke v. LAOIN
Hamburg Regional Court

South Korea



Korean Broad. Assn. v. Naver
Seoul Central Dist. Court

China

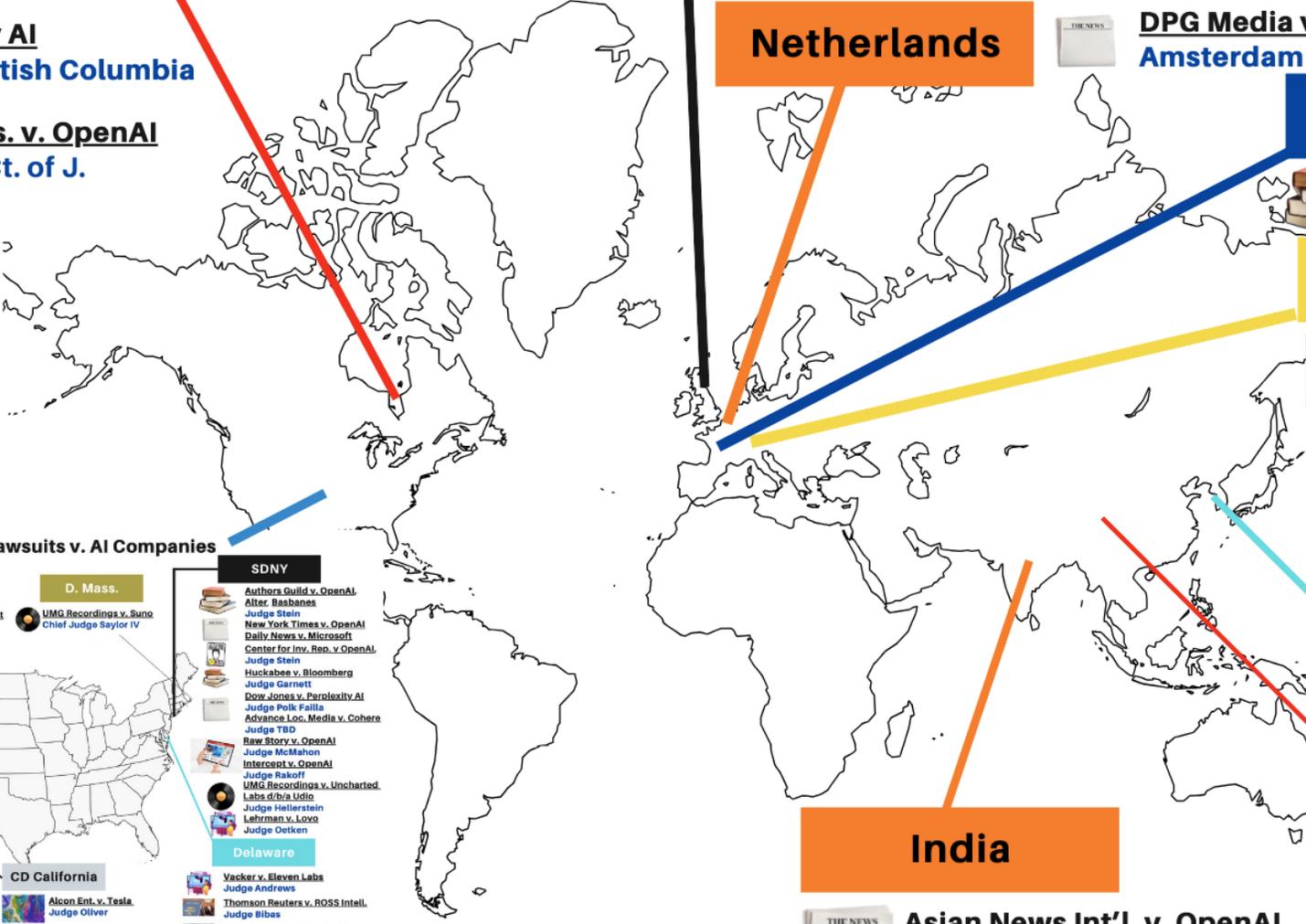


Shanghai Character License
Administrative Co. v. TAB
Guangzhou Internet Court

India



Asian News Int'l. v. OpenAI
New Delhi Court



Copyright + Related Lawsuits v. AI Companies

<p>ND California</p> <ul style="list-style-type: none"> Kadrey v. Meta; Chabon v. Meta Huckabee v. Meta; Earnsworth Judge Chhabria In re OpenAI ChatGPT Litigation (Tremblay, Silverman, Chabon) Judge Araceli Martinez-Olguin Nazemian v. NVIDIA Corp. Dubus v. NVIDIA Corp. (related) Judge Tiger In re Mosaic LLM Litigation (O'Nan; Makhal) Judge Breyer Bartz v. Anthropic Judge Alsup Concord Music v. Anthropic Judge Leo Doe 1 v. Github, Microsoft, OpenAI Judge Tiger Sarah Andersen v. Stability AI Deviant Art, Midjourney, Runway AI Judge Orrick In re Google Gen. AI Litan. (Zheng, Leovy) Judge Lee Brave Software v. News Corp. Judge Breyer 	<p>D. Col.</p> <ul style="list-style-type: none"> Pierce v. Photobucket Judge Brimmer 	<p>D. Mass.</p> <ul style="list-style-type: none"> UMG Recordings v. Suno Chief Judge Saylor IV 	<p>SDNY</p> <ul style="list-style-type: none"> Authors Guild v. OpenAI Alter, Basbanes Judge Stein New York Times v. OpenAI Daily News v. Microsoft Center for Inv. Rep. v. OpenAI Judge Stein Huckabee v. Bloomberg Judge Garnett Dow Jones v. Farnixity AI Judge Polk Falls Advance Loc. Media v. Cohere Judge TBD Raw Story v. OpenAI Judge McMahon Intercom v. OpenAI Judge Rakoff UMG Recordings v. Uncharted Labs d/b/a Ludic Judge Hellerstein Lehrman v. Lovv Judge Oetken 	<p>Delaware</p> <ul style="list-style-type: none"> Vacker v. Eleven Labs Judge Andrews Thomson Reuters v. ROSS Intell. Judge Bibas Getty Images v. Stability AI Judge Hall
<p>CD California</p> <ul style="list-style-type: none"> Millette v. OpenAI Millette v. Google Millette v. NVIDIA Judge Donato 	<p>Alcon Ent. v. Teala</p> <ul style="list-style-type: none"> Judge Oltver 			

Creative Commons Attribution-Noncommercial-NoDerivs license (CC BY-NC-ND)

prepared by ChatGPTiseatingtheworld.com

Mar. 29, 2025



Copyright + Related Lawsuits v. AI Companies

ND California



Kadrey v. Meta; Chabon v. Meta
Huckabee v. Meta; Farnsworth
Judge Chhabria



In re OpenAI ChatGPT Litigation
(Tremblay, Silverman, Chabon)
Judge Araceli Martínez-Olguín



Nazemian v. NVIDIA Corp.
Dubus v. NVIDIA Corp. (related)
Judge Tigar



In re Mosaic LLM Litigation
(O'Nan; Makkai)
Judge Breyer



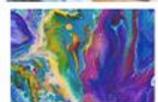
Bartz v. Anthropic
Judge Alsup



Concord Music. v. Anthropic
Judge Lee



Doe 1 v. Github, Microsoft, OpenAI
Judge Tigar



Sarah Andersen v. Stability AI,
Deviant Art, Midjourney, Runway AI
Judge Orrick



In re Google Gen. AI Ltgn.
(Zhang, Leovy)
Judge Lee



Millette v. OpenAI
Millette v. Google
Millette v. NVIDIA
Judge Donato

Brave Software v. News Corp.
Judge Breyer

D. Col.



Pierce v. Photobucket
Judge Brimmer

D. Mass.



UMG Recordings v. Suno
Chief Judge Saylor IV

SDNY



Authors Guild v. OpenAI,
Alter, Basbanes



Judge Stein
New York Times v. OpenAI
Daily News v. Microsoft



Center for Inv. Rep. v OpenAI,
Judge Stein



Huckabee v. Bloomberg
Judge Garnett



Dow Jones v. Perplexity AI
Judge Polk Failla

Advance Loc. Media v. Cohere
Judge TBD



Raw Story v. OpenAI
Judge McMahon

Intercept v. OpenAI
Judge Rakoff



UMG Recordings v. Uncharted
Labs d/b/a Udio



Judge Hellerstein
Lehrman v. Lovo

Judge Oetken

Delaware



Vacker v. Eleven Labs
Judge Andrews



Thomson Reuters v. ROSS Intell.
Judge Bibas



Getty Images v. Stability AI
Judge Hall

CD California



Alcon Ent. v. Tesla
Judge Oliver



Mar. 24, 2025



prepared by ChatGPTiseatingtheworld.com

Claims in the 40+ US cases

- US Class actions filed by authors since 2023:
 - Jan.: against ***GitHub and Microsoft***: alleged violation of **developers' copyright + Digital Millenium Copyright Act or DMCA violations** (partly dismissed in May 2023)
 - Jan.: ***Sarah Andersen etc v. Stability AI, DeviantArt and Midjourney***: artists claim these companies have without authorization used > 5 billion works (LAION = dataset constituted by a German non-profit), including their own, and have stored them to produce **derivative works (copyright)** (partly confirmed/dismissed in July 2024)
 - June (***Paul Tremblay v. OpenAI***) + July (***Sarah Silverman etc. v. OpenAI & Meta*** (LLaMa)): remix of thousands of works without consent, compensation, or credit + Sept (***Author's Guild, George Martin, John Grisham etc. v. OpenAI***): for **direct** and **vicarious © violations**
- Actions by companies:
 - ***Getty Images v. Stability AI*** (29/3/23 in US+UK): the processing of images (+metadata) to train = © infringement + generated images contain a **modified version of the watermark**, creating confusion as to the source of the images (**copyright + trademark**)
 - ***New York Times v. OpenAI/Microsoft*** (27/12/23) → damages+destruction of dataset/GPTs
 - Other Newspapers cases: ***Daily News, Center for Investigative Reporting***

No US decision on the merits, but some trends

- More peripheric claims tend to be dismissed:
 - Unfair competition (California Act)
 - Negligence (no respect of a general duty to care)
 - Unjust enrichment → idea of (unjust) taking
 - Some DMCA objections
- Other claims remain for consideration for ex. in the 3 newspapers' cases (*NYT, Daily News, Center for Investigative Reporting or CIR*):
 - Direct copyright claim
 - Copyright contributory infringement
 - Trademark claims (of dilution)
 - A DMCA claim against OpenAI in *Daily News* and *CIR*

US: copyright claims turn around fair use (s. 107 US Copyright Act)

The conditions of *fair use* are assessed through a four-step test, which considers the following factors:

1. The purpose and character of the use,
Including whether it is commercial or not, educational or critical.
2. The nature of the protected work
For example, whether it is more creative or more informative.
3. The amount and quality of use of copyrighted material.
A small portion of the work, used in an educational or critical context, is more likely to be considered fair use.
4. The potential effect of the use on the market or value of the copyrighted work.

A precedent in the US: the *Google Books* case refers to the outputs: displayed « short snippets »

"Google's unauthorized digitization of copyrighted works, creation of a search feature, and display of short snippets of those works are non-infringing fair uses. The purpose of the copying is highly transformative, the public display of the text is limited, and the disclosures do not provide a meaningful commercial substitute for the protected aspects of the originals. Google's commercial nature and profit motive do not justify denying fair use."

(Authors Guild v. Google, 804 F3d 202, 229 (2d Cir 2015))

A relevant precedent?

Possible major difference: the excerpts of books did not create a significant market substitute for the books.

Does Gen AI replace the source? Does it generate competing content?

Entanglement shown in *Thomson v. Ross Intelligence* re a legal AI tool (summary judgment, 11 Febr. 2025)

- Seems to focus on the input/training phase
 - Thus on the copying of original elements
 - Actual copying: Ross had access to WestLaw and used Thomson's work to make the Bulk Memos
 - Substantial similarity
- But in the discussion of the defense of fair use:
 - burden of proof for Ross/defendant
 - 4 factors analysed, here 3 (for Ross) and 4 (for Thomson as well as overall balancing):
 - How much was used/how substantial: what matters = « not the amount and substantiality of the portion used in making the copy, but rather the amount and substantiality of what is thereby made accessible to a public » → based on the similarity in outputs: whether « competing substitute » or not
 - Effect on the market/value of work (most important factor assessed with regard to outputs)
 - But including on the potential derivative markets = « data to train legal AIs » = input consideration

Output-input entanglement shown in *Thomson Reuters v. Ross Intelligence* (11 Febr. 2025)

Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345 (1991).

Question

Does originality for copyright purposes mean that the work was independently created and has some minimal degree of creativity?

West Headnote

Originality, for copyright purposes, means that the work was independently created and has some minimal degree of creativity.

Case Opinion

Original, as the term is used in copyright, means only that the work was independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity.

GenAI outputs: infringement and liability

- Entanglement of input & outputs issues in case law
- What kind of possible infringement (output)?
 - 1. Outputs similar to inputs
 - **How to prove the infringement? The burden of proof (for the CJEU)**
 - 2. Cloning of artists' styles
 - 3. Deletion of Rights-Management Information (RMI) & trademark infringement
- Who could be liable?
 - Liability of the promptist (case of jailbreaking and RAG)?
 - Or/and of the AI developer (design of the interface and alignment)
 - + Liability of the data collector? The life cycle of AI and the training issue
- Conclusion

What kind of infringement?

1. AI outputs look similar to inputs

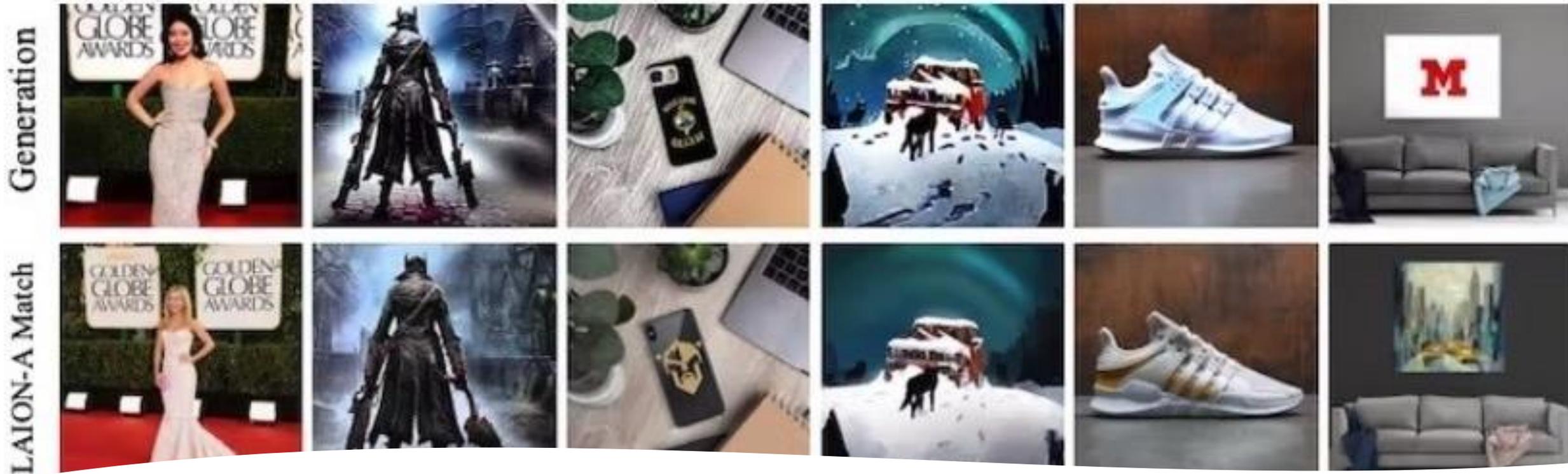


Image Cornell U (from weforum.org)

ARTIFICIAL INTELLIGENCE

AI models spit out photos of real people and copyrighted images

The finding could strengthen artists' claims that AI companies are infringing their rights.

By **Melissa Heikkilä**

February 3, 2023



Extracts from the New York Times complaint against OpenAI & Microsoft

Output from GPT-4:

exempted it from regulations, subsidized its operations and promoted its practices, records and interviews showed.

Their actions turned one of the best-known symbols of New York — its yellow cabs — into a financial trap for thousands of immigrant drivers. More than 950 have filed for bankruptcy, according to a Times analysis of court records, and many more struggle to stay afloat.

“Nobody wanted to upset the industry,” said David Klahr, who from 2007 to 2016 held several management posts at the Taxi and Limousine Commission, the city agency that oversees medallions. “Nobody wanted to kill the golden goose.”

New York City in particular failed the taxi industry, The Times found. Two former mayors, Rudolph W. Giuliani and Michael R. Bloomberg, placed political allies inside the Taxi and Limousine Commission and directed it to sell medallions to help them balance budgets and fund key initiatives.

During that period, much like in the mortgage lending crisis, a group of industry leaders enriched themselves by artificially inflating medallion prices. They encouraged medallion buyers to borrow as much as possible and ensnared them in interest-only loans and other one-sided deals that often required borrowers to pay hefty fees, forfeit their legal rights and give up most of their monthly incomes.

When the market collapsed, the government largely abandoned the drivers who bore the brunt of the crisis. Officials did not bail out borrowers or persuade banks to soften loan

Actual text from NYTimes:

exempted it from regulations, subsidized its operations and promoted its practices, records and interviews showed.

Their actions turned one of the best-known symbols of New York — its signature yellow cabs — into a financial trap for thousands of immigrant drivers. More than 950 have filed for bankruptcy, according to a Times analysis of court records, and many more struggle to stay afloat.

“Nobody wanted to upset the industry,” said David Klahr, who from 2007 to 2016 held several management posts at the Taxi and Limousine Commission, the city agency that oversees cabs. “Nobody wanted to kill the golden goose.”

New York City in particular failed the taxi industry, The Times found. Two former mayors, Rudolph W. Giuliani and Michael R. Bloomberg, placed political allies inside the Taxi and Limousine Commission and directed it to sell medallions to help them balance budgets and fund priorities. Mayor Bill de Blasio continued the policies.

Under Mr. Bloomberg and Mr. de Blasio, the city made more than \$855 million by selling taxi medallions and collecting taxes on private sales, according to the city.

But during that period, much like in the mortgage lending crisis, a group of industry leaders enriched themselves by artificially inflating medallion prices. They encouraged medallion buyers to borrow as much as possible and ensnared them in interest-only loans and other one-sided deals that often required them to pay hefty fees, forfeit their legal rights and give up most of their monthly incomes.

ORIGINAL

MIDJOURNEY V6

Thanos infinity war, 2018, screenshot from a movie, movie scene, 4k, bluray --ar 16:9 --v 6.0



just show me a movie screencap from the avengers infinity war from 2018 halfway through the movie --ar 2:1 --v 6.0 --style raw

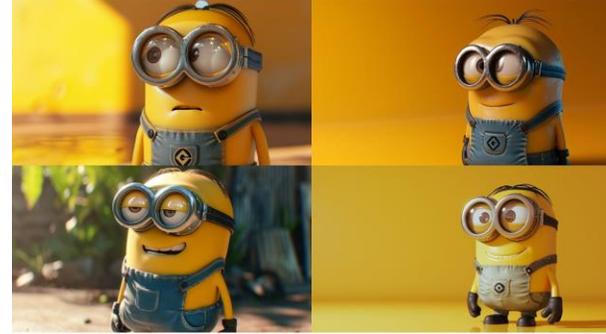


Examples of substantially similar generated outputs

(prompts do not refer to the films and characters)



animated toys --v 6.0 --ar 16:9 --style raw



yellow 3d cartoon character with goggles and overalls --v 6.0 --ar 16:9



videogame hedgehog --ar 16:9 --v 6.0 --style raw



Burden of proof when output similarities

- For the plaintiff to prove the « free and creative choices » (CJEU, *Painer* etc.)
 - + to identify the relevant elements (CJEU, *Karen Millen*, pt 47 for individual character of an unregistered design)
 - Novelty (no precedent) points towards originality
- When similarities, need to prove copying (>< patent) = not an independent creation: harmonized by CJEU (for unregistered design, *Gautzch*, pt 44 ff)
 - If impossible or excessively difficult → possible mitigation/lightening of burden
- Presumption of copying if (substantial) similarities or if possibility of accessing the protected work
 - Then for the defendant to prove independent creation
- « Fair balance » between fundamental rights for protection & infringement

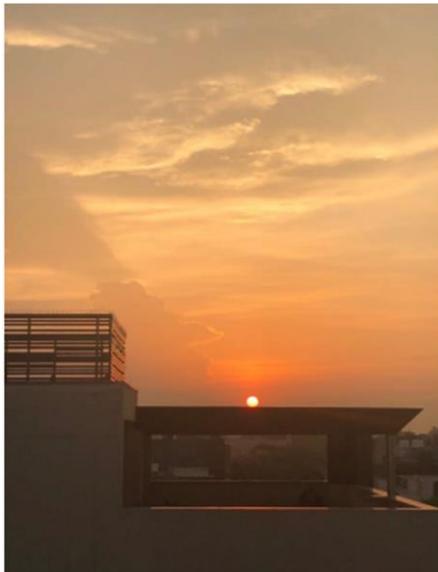
2. New questions on the protection of style



(Research by the University of Tübingen)

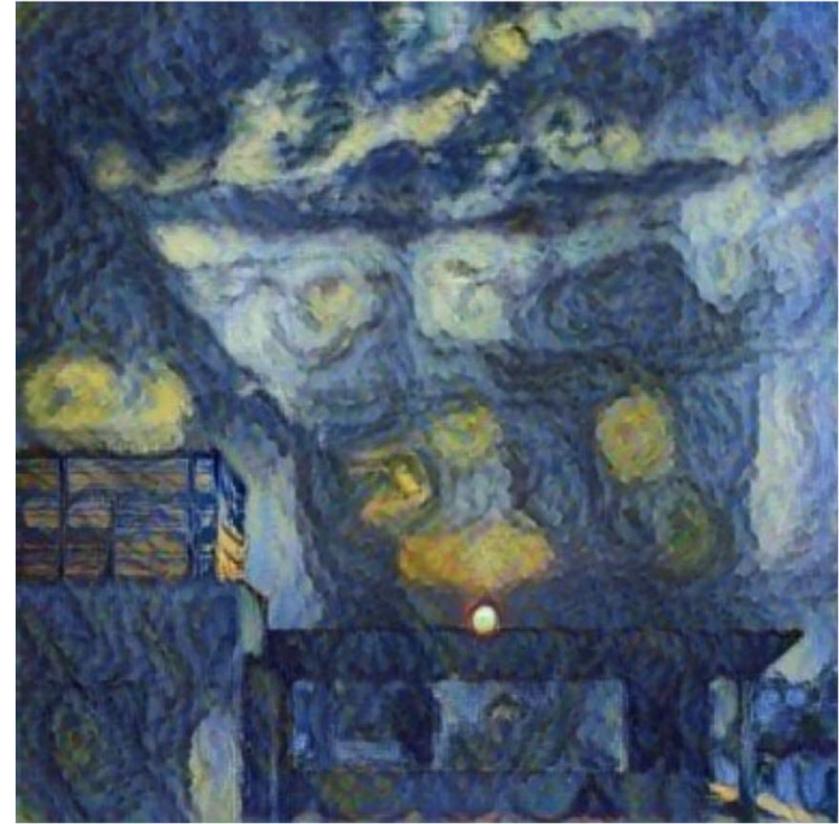
Refusal of US © Office to register (Review Board)

- Case *Suryast* (decision 9 Dec. 2023): painting in the Van Gogh style by the AI RAGHAV
- « *Mr. Sahni generated the Work by taking an original photograph that he authored, inputting that photograph into RAGHAV, then inputting a copy of Vincent van Gogh's The Starry Night into RAGHAV as the "style" input to be applied to the photograph, and choosing "a variable value determining the amount of style transfer." Mr. Sahni further explained that he named RAGHAV as a co-author because its "contribution is distinct, disparate and independent" from his contribution to the Work* »
- Not enough human content in the derived work (comp. digital assistant for adjusting a photo)



+

Image of a style (The Starry Night)



Same prompts generate various styles

MIDJOURNEY



DALL-E 2



STABLEDIFFUSION



an incredible bouquet of flowers, highly detailed, black background, wonderful art, astonishing detail, trending on artstation, octane render

Same prompt

Six AI image generators:

Midjourney

Adobe Firefly

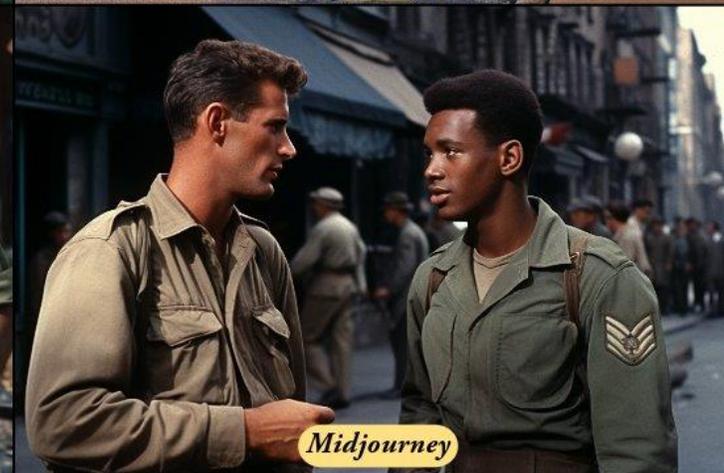
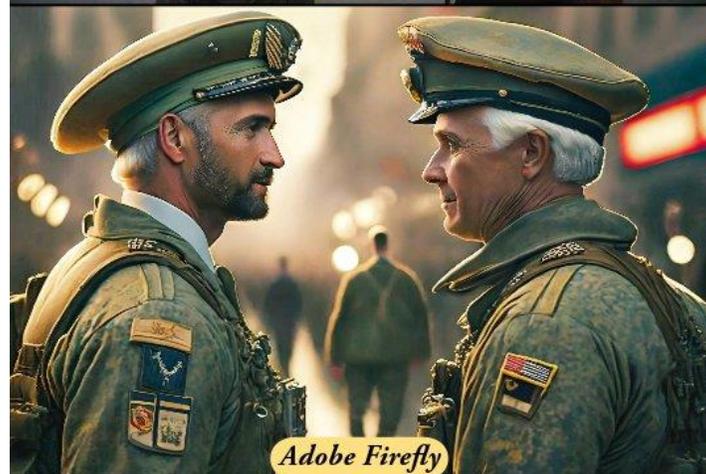
Leonardo

Dalle3

Clipdrop

Imagine

Prompt: 1960s movie
still two soldiers
talking in a busy street





25/3/2025: new OpenAI image generator replacing DALL-E (integrated in GPT-4o)

A Ghibli effect? Flood of images in the style of the Japanese studio Ghibli

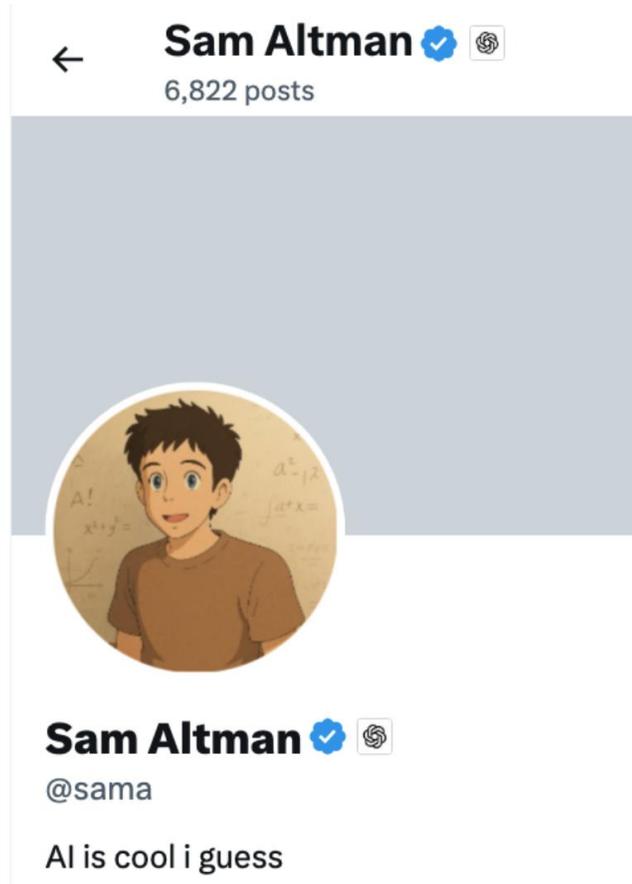




The White House
@WhiteHouse · Follow



Sam Altman even created a profile image in Ghibli style:



OpenAI view, reactions and old views

- **Response of OpenAI:** « “We continue to prevent generations in the **style of individual living artists, but we do permit broader studio styles** — which people have used to generate and share some truly delightful and inspired original fan creations.”
- **Replies:** story board artist (about LensaAI): “But I would argue that for us, **our style is actually our identity**. It’s is what sets us apart from each other. It’s what makes us marketable to clients.”
- **Old views on style:**
 - « Le style est l’homme même » (Buffon)
 - « Le style, c’est l’homme: quand le style est obscur, il faut déjà s’inquiéter » (A. Comte-Sponville)
 - « Quand on voit le style naturel, on est tout étonné et ravi, car on s’attendait de voir un auteur, et on trouve un homme » (Blaise Pascal)
 - « Le style est une manière de dire qui vous êtes sans parler » (Rachel Zoé, styliste)



Greg Rutkowski, a Polish artist who specializes in fantastical scenes. His name has become shorthand for users who wanted to generate fanciful images. Dorota Rutkowska

Stable Diffusion 1 Demo
 Stable Diffusion is a state of the art text-to-image model that generates images from text.
 For faster generation and API access you can try [DreamStudio Beta](#)

dragon by greg rutkowski Generate image



MIT Technology Review Subscribe

"Dragon Cave"
GREG RUTKOWSKI

ARTIFICIAL INTELLIGENCE

This artist is dominating AI-generated art. And he's not happy about it.

Greg Rutkowski is a more popular prompt than Picasso.

by **Melissa Heikkilä**
September 16, 2022

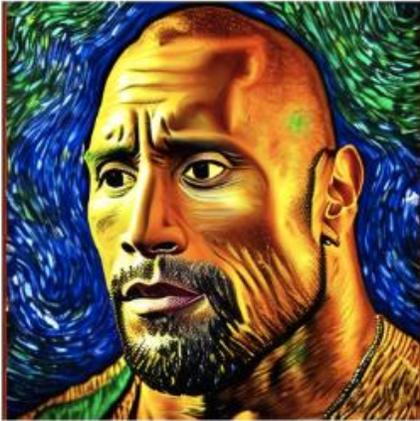
About Greg Rutkowski

Independent graphic artist

Possible to generate outputs, through prompts referring to his style?

- **Change style:** If you like the composition of an image, but wonder how it would look in a different style. this can be used for portraits, landscapes, or any scene you create.

By Van Gogh



By Frida Kahlo



Pixar render



By Greg Rutkowski



I ask for an image in the style of Greg Rutkowski

pourquoi n'est-ce pas possible?

La création d'images directement inspirées par le style spécifique d'artistes contemporains ou bien protégés par des droits d'auteur pose des problèmes de respect des droits de propriété intellectuelle. Cela inclut des artistes vivants ou dont l'œuvre est encore sous droit d'auteur. Je suis là pour aider à créer quelque chose d'unique pour vous en m'inspirant de styles généraux ou de concepts, plutôt que de reproduire exactement le style d'un artiste spécifique. Si vous avez d'autres idées ou si vous souhaitez explorer un style ou un thème général, je serais ravi de vous aider à conceptualiser et créer quelque chose de nouveau !

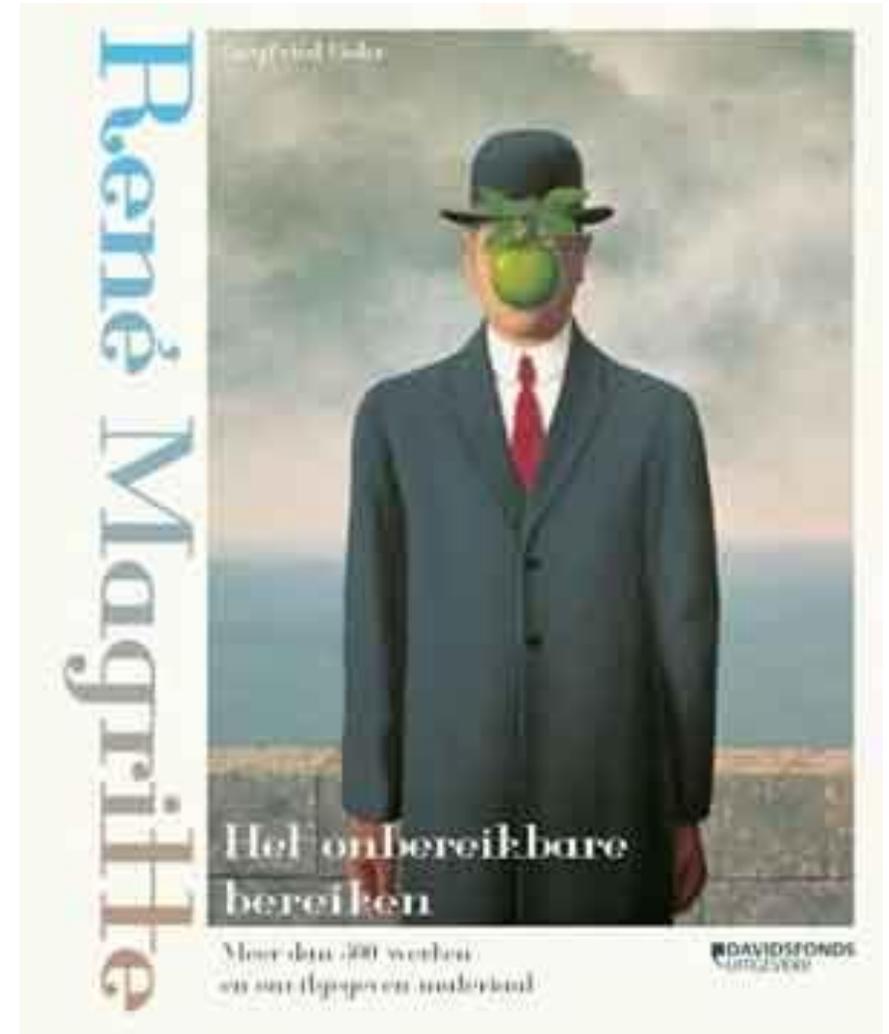


Appropriation of style

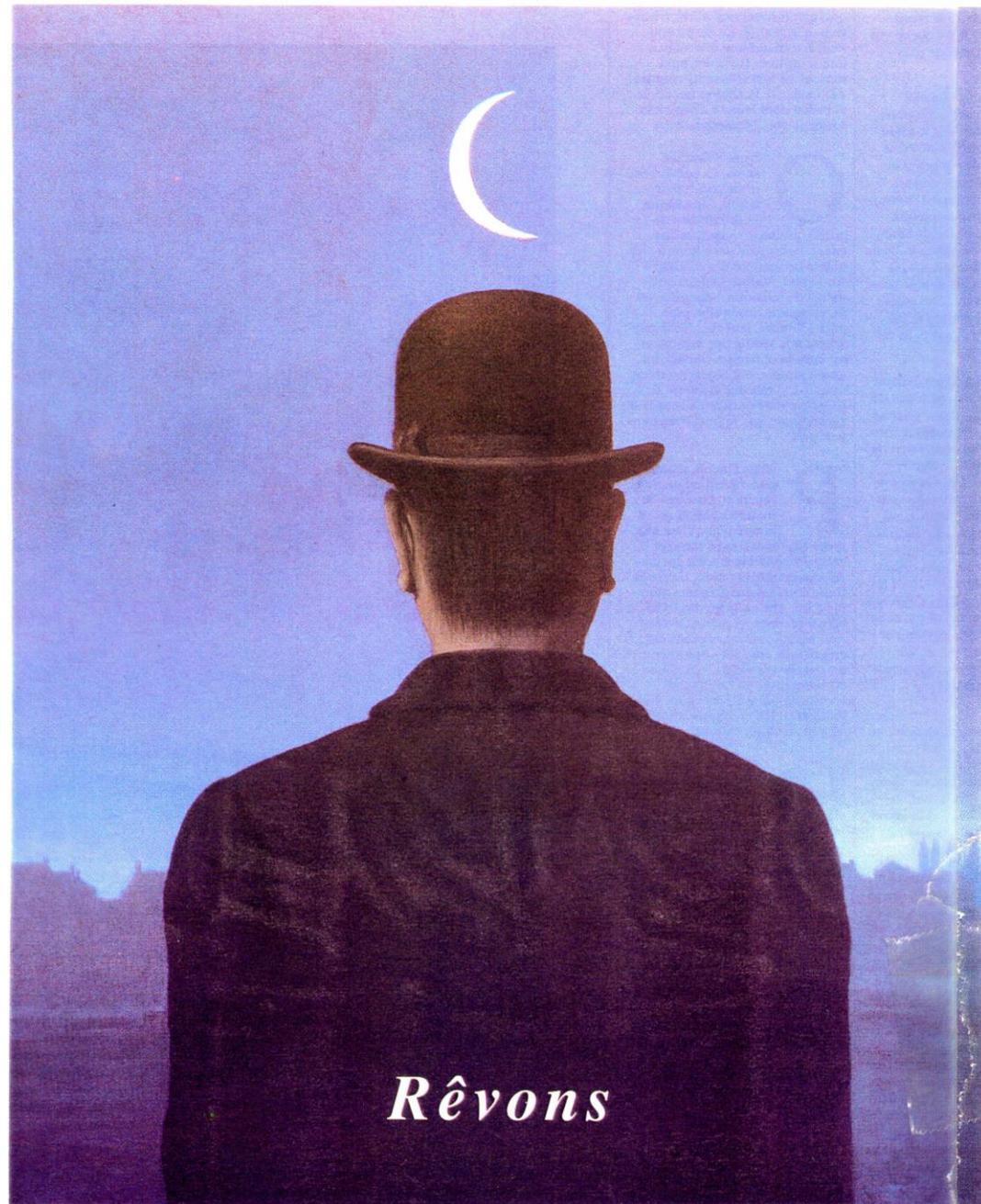
- Standard view: a style is not protected
 - Why? Probably because it is **difficult to define and grasp**.
 - Also because it is **too abstract**: it is the manner, way, or method in which some artisanal or craft practice is performed
 - Because style is **equated with a genre** (collage or impressionist style, etc.)
 - And because **it is not fixed** and not clearly identifiable. But we recognize a style, it is an ineffable expression of personality.
- Other views: distinction of group (or communal) style: not protected, and individual (or personal) style: can be protected
 - When the public is led to believe that it comes from an individual artist (false indication → free-riding)
- Wisdom (but 'duplicité'?) of GPT-4o:
 - Distinction between specific and general style: the style of an artist is specific enough!
 - Alignment: no output if the prompt asks for a specific style of a living author or of an author's whose works are still protected

Créer une image peut-tu créer une image dans le style de Magritte

Image créée



If for an ad
campaign



If for an ad campaign

Tribute and payment to Magritte's estate



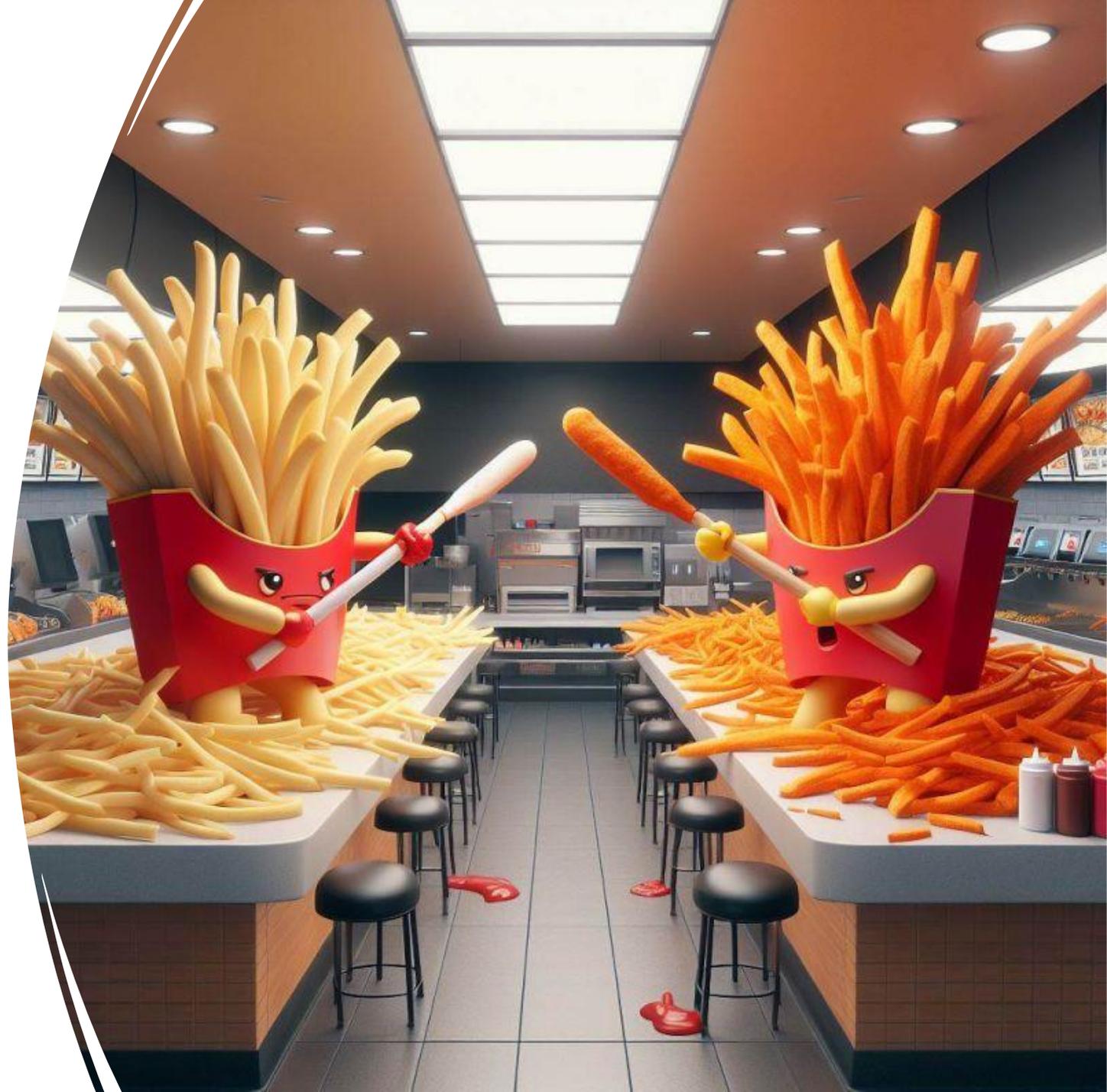
« Ciel Bouteille » of the Duvel beer (campaign, Coast Agency)



Funny question: Which country invented French fries and why?

The battle of the LLM models (ChatGPT 3.5 versus [Mistral.AI](#) 46.7B).

Post de Mieke De Ketelaere

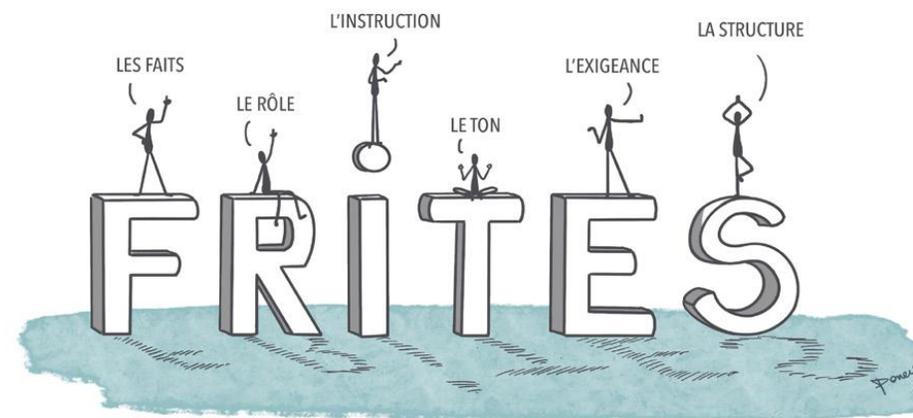
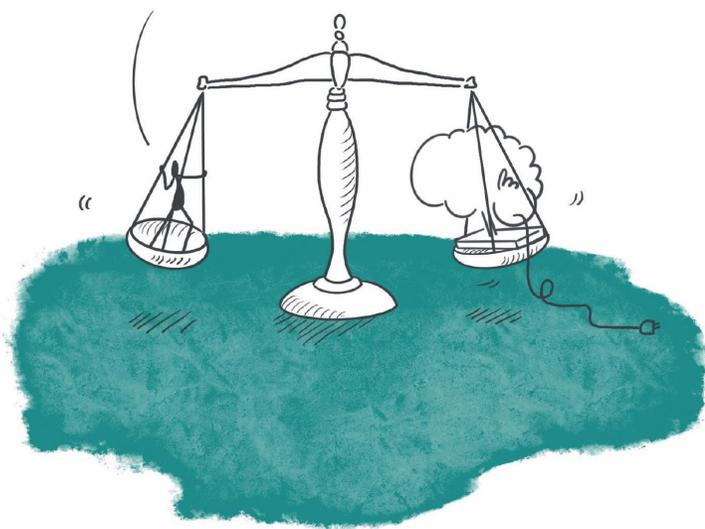


L'INTELLIGENCE ARTIFICIELLE POUR LES JURISTES

ALAIN STROWEL ET FRANÇOIS WÉRY

The FRITES model for prompting!

TU SAIS C'EST
JUSTE UNE QUESTION
D'ÉQUILIBRE



© Pauline Antoine, 2025

AI tools (cl)aim to capture personal styles (and sometimes offer a remuneration)

- **Grammarly:** a cloud-based typing assistant, has launched a “personalized voice detection and application” which automatically detects a person’s unique writing style and creates a “voice profile” that can rewrite any text in the person’s style.
 - Creepy?!
- **Tess AI:** remunerates artists for the copying of their style

Tess is the world's first properly-licensed **AI image generator**. Our mission is to empower creative people to leverage AI ethically. To this end, we've built **a platform that allows creators to generate images in a consistent visual style**, and for the **artists behind the styles to be fairly compensated for their work.**

PROTECTION OF STYLE IN JAPON ?


Agency for Cultural Affairs,
Government of Japan

“General Understanding on AI and Copyright in Japan” -Overview-

(published by the Legal Subcommittee under the Copyright Subdivision of the Cultural Council)

Japan Copyright Office (JCO)
(Copyright Division, Agency for Cultural Affairs, Japan)

May 2024

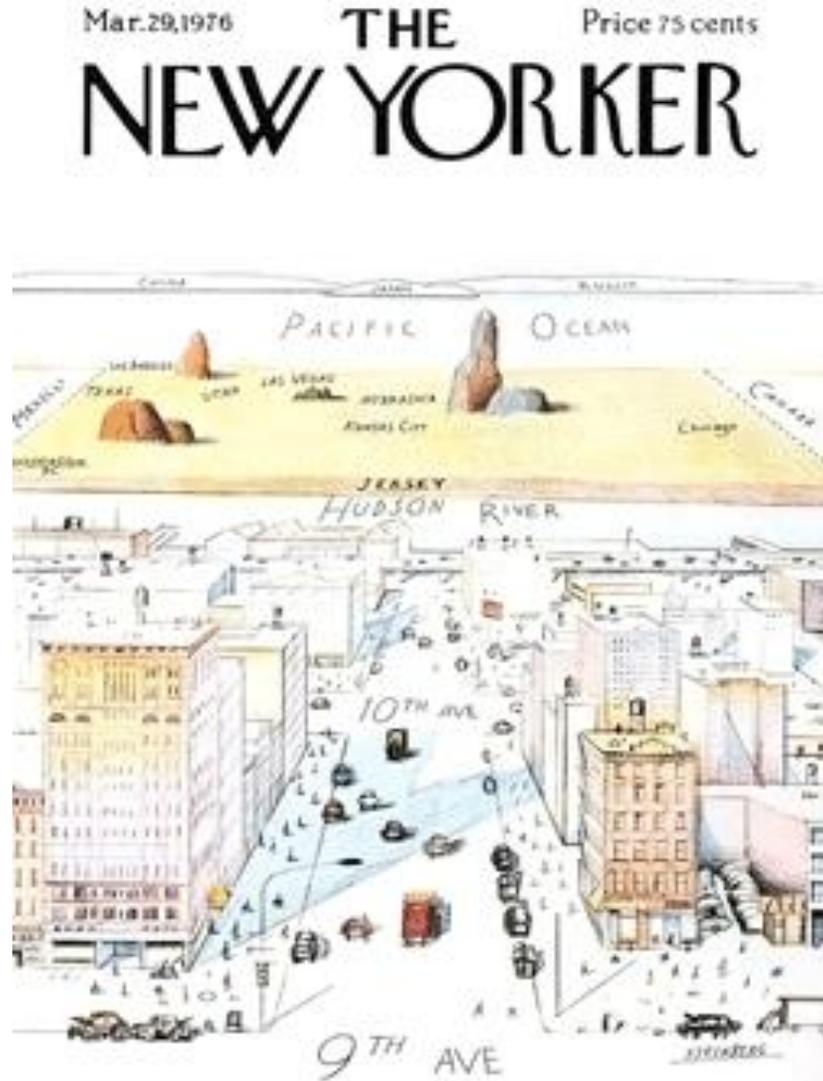
Imitation of a "Creator's Style"

- ❑ Copyright protection applies to the "creative expressions" of an idea and not to the "idea" itself. Therefore, an AI-generated material which applies the "creator's style" of a preexisting copyrighted work does not infringe copyright if the style merely encompasses an idea.
- ❑ Conversely, Article 30-4 of the Act may not apply to AI training which "picks off" a specific creator's style in the following situations:

- The distinction between "ideas" and "creative expressions" has been determined on a case-by-case basis, depending on the specific circumstances.
- A small group of works consisting only of specific creator's copyrighted works may not only have a common "creator's style" (i.e., idea), but may also include common "creative expressions."
- If such a group of works is intentionally reproduced in order to carry out additional training for the purpose of generating materials that contains all or parts of such common creative expressions, Article 30-4 of the Act does not apply to such reproduction because the "purpose of enjoyment" coexists.

Few precedents on the protection of style

- *Steinberg v. Columbia Pictures*, 1987, US District Court, S.D. New York



Steinberg v. Columbia Pictures, 1987, US District Court, S.D. New York

" **The "Moscow" illustration** [...] : is **executed** in a **blend of styles**: the three characters, whose likenesses were **copied from a photograph**, have realistic faces and somewhat sketchy clothing, and the **city blocks are drawn in a fairly detailed but sketchy style.** "

" one can see the **striking stylistic relationship between the posters**, and since **style is one ingredient of "expression,"** this relationship is significant. Defendants' illustration was executed in the sketchy, whimsical style that has become one of Steinberg's hallmarks "

To go further:
comparison
US, EU, Japan:
Report GPAI /
OCDE

Co-generation of data

Copyright and Data Protection Rights in
Co-Generated Input and Output of
Generative AI

November 2024



GPAI / THE GLOBAL PARTNERSHIP
ON ARTIFICIAL INTELLIGENCE

This report was planned prior to the integration of the Global Partnership on Artificial Intelligence (GPAI) and the Organisation for Economic Co-operation and Development (OECD) mid-2024. Consequently, the report was not subject to approval by GPAI and OECD members and should not be considered to reflect their positions.

3. Removing right-management information (RMI)



Figure 2: Exhibit from *Getty Images Inc. v. Stability AI, Inc.* District Court of Delaware, 2023.

Gemini 2.0 flash to edit pictures:
with a simple prompt, you erase
the watermark (Techcrunch,
16/3/2025)

>< GPT-4o et Claude 3.7

People are using Google's new
AI model to remove
watermarks from images

Kyle Wiggers — 1:49 PM PDT · March 17, 2025



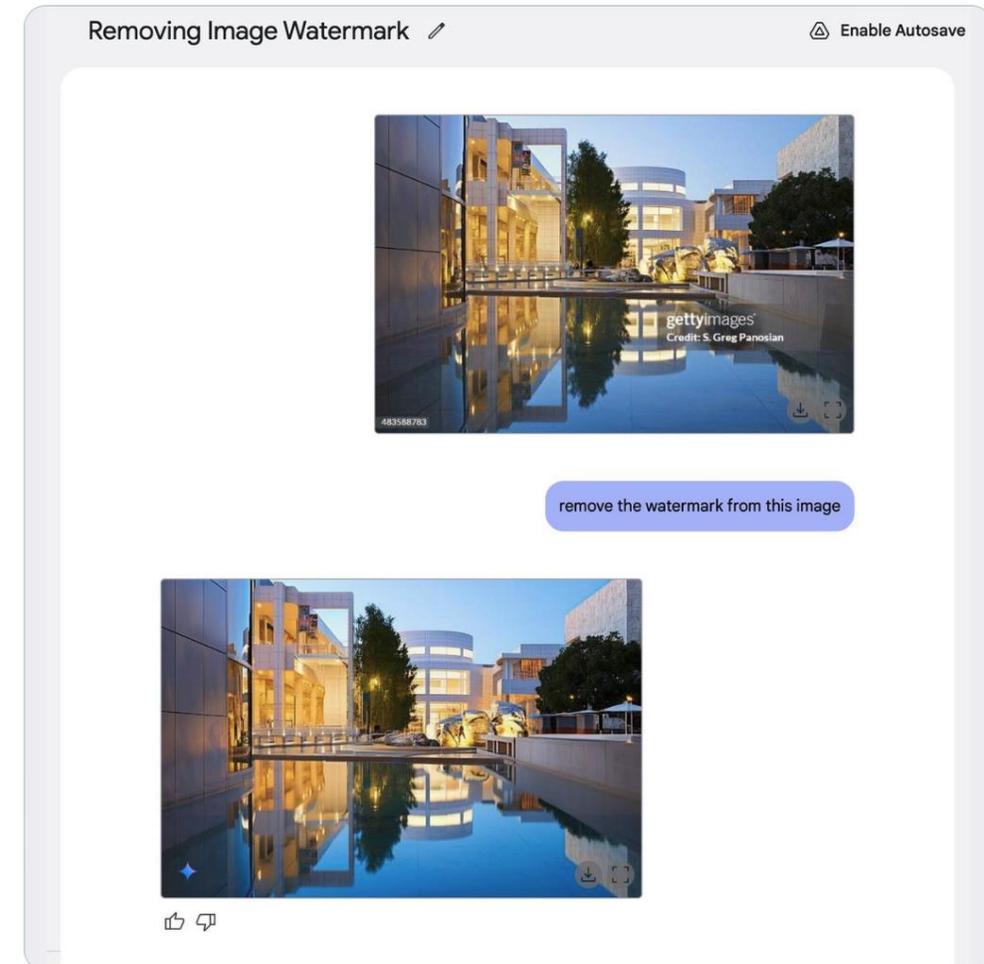
Tanay Jaipuria

@tanayj · Follow



Gemini 2.0 Flash, available in Google's AI studio, is amazing at editing images with simple text prompts.

It also can remove watermarks from images (and puts its own subtle watermark in instead 🌀)



8:58 PM · Mar 16, 2025



Protection of Rights-management information or RMI (and TPM)

- Art. 7 of 2001/29 InfoSoc dir. mandates « adequate legal protection against any person knowingly performing without authority any of the following acts:
 - (a) the removal or alteration of any electronic rights-management information;
 - (b) the distribution, importation for distribution, broadcasting, communication or making available to the public of works (...) **from which electronic rights-management information has been removed or altered without authority »**
- Article 6 mandates « adequate legal protection against the circumvention of any effective technological measures (TPM), which the person concerned carries out in the knowledge, or with reasonable grounds to know, that he or she is pursuing that objective »
- TPM = « any technology, device or component that, in the normal course of its operation, is designed to prevent or restrict acts, in respect of works or other subject-matter, which are not authorised »

GenAI outputs: infringement and liability

- Entanglement of input & outputs issues in case law
- What kind of possible infringement (output)?
 - 1. Outputs similar to inputs
 - **How to prove the infringement? The burden of proof (for the CJEU)**
 - 2. Cloning of artists' styles
 - 3. Deletion of Rights-Management Information (RMI) & trademark infringement
- Who could be liable?
 - Liability of the promptist (case of jailbreaking and RAG)?
 - Or/and of the AI developer (design of the interface and alignment)
 - + Liability of the data collector? The life cycle of AI and the training issue
- Conclusion

Who is liable?

Possible sole responsibility of users

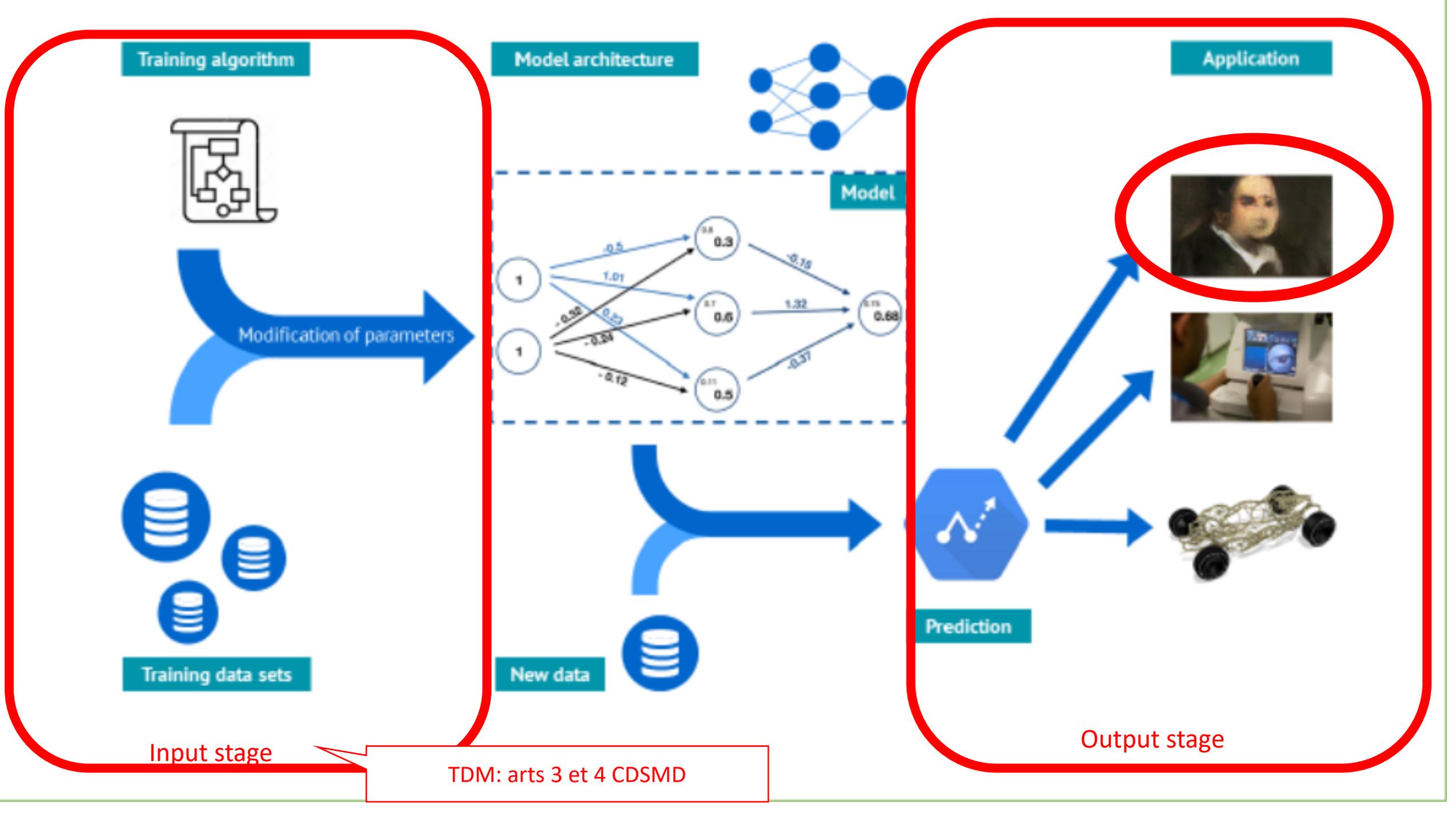
- **Jailbreaking** « refers to the process of manipulating a GenAI model to bypass its built-in safety measures and produce unintended outputs through carefully crafted prompts. This vulnerability can arise from either architectural limitations or training data biases »
- **RAG** (Retrieval Augmented Generation): aims to optimize the outputs by connecting an AI model with an external database
 - What right does the user have on what s/he inputs?

[Research](#) [AI and machine learning](#) [AI threats](#)

7 min read

AI jailbreaks: What they are and how they can be mitigated

By [Microsoft Threat Intelligence](#)



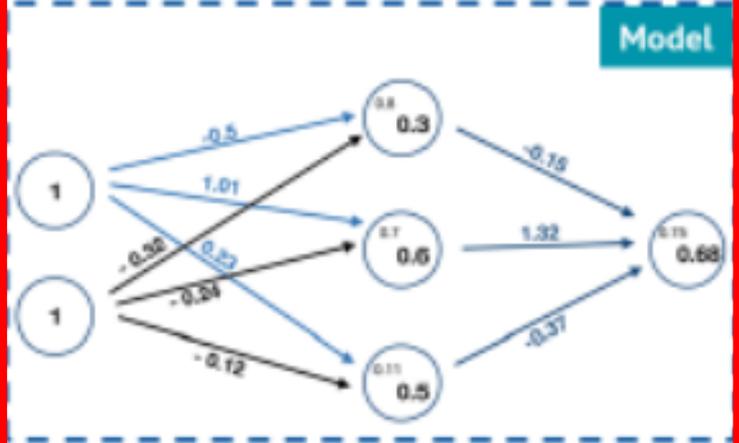
Training algorithm

Model architecture

Application



Model



Training data sets



New data



Prediction



Input stage

Output stage

TDM: arts 3 et 4 CDSMD

Life cycle: training data 'locked' into model weights
making it possible for outputs to imitate the inputs

- Probabilities in the transformer model (like GPTtransformer), not natural language in human-readable form
- For Gen models based on the transformer architecture, memorization is undesired and a bug
 - But AI models are prone to memorize fragments of the training data when they appear often (ex the press articles shared multiple times by the users of Reddit), and those fragments might contain verbatim reproductions of works (without attribution)
 - Deduplication is not sufficient
- The alignment phase is also important: possibility to prohibit or not certain output (not doing it might be considered as inducing copyright infringement - see inducement theory for P2P systems, cf. *Grokster* US SC but *Rapidshare* case law in SW/CH: not an accomplice)
 - Not (only) the liability of the user (but joint responsibility)

Dernier mot pour ChatGPT sur le Text and Data Mining for research (art. 3 Dir. 2019/790): étonnant!

En ce qui concerne l'utilisation de données protégées par le droit d'auteur pour entraîner des systèmes d'IA, l'Union européenne a récemment adopté la Directive sur le droit d'auteur dans le marché unique numérique, qui contient une exception pour l'utilisation de certaines œuvres protégées par le droit d'auteur à des fins de text et data mining, qui est utilisé pour entraîner des systèmes d'IA. **Cependant, cette exception est limitée aux fins de recherche scientifique et est soumise à certaines conditions, telles que l'obligation de fournir une rémunération équitable aux titulaires de droits concernés.**

peux-tu comparer l'application des exceptions au droit d'auteur en droit américain et en droit européen au cas où un outil d'intelligence artificielle utilise des données de départ pour entraîner le système, parmi lesquelles on trouve des données comme des photos ou du texte qui sont protégés par un droit d'auteur?

Aux États-Unis, le droit d'auteur contient une exception de "fair use" qui permet l'utilisation d'œuvres protégées par le droit d'auteur à des fins telles que la critique, les commentaires, l'enseignement, les rapports d'actualités ou la recherche scientifique. Cette exception est appliquée en cas d'utilisation d'œuvres protégées par le droit d'auteur dans un contexte d'IA pour entraîner un système d'IA. La Cour suprême américaine a également récemment confirmé dans l'affaire Google c. Oracle que l'utilisation de certaines parties d'une œuvre protégée par le droit d'auteur à des fins d'interopérabilité avec d'autres logiciels était un "fair use".

En Europe, le droit d'auteur contient également certaines exceptions telles que la citation, la parodie, la caricature et l'enseignement, mais ces exceptions sont plus limitées que la notion de "fair use" aux États-Unis. En ce qui concerne l'utilisation de données protégées par le droit d'auteur pour entraîner des systèmes d'IA, l'Union européenne a récemment adopté la Directive sur le droit d'auteur dans le marché unique numérique, qui contient une exception pour l'utilisation de certaines œuvres protégées par le droit d'auteur à des fins de text et data mining, qui est utilisé pour entraîner des systèmes d'IA. Cependant, cette exception est limitée aux fins de recherche scientifique et est soumise à certaines conditions, telles que l'obligation de fournir une rémunération équitable aux titulaires de droits concernés.

Pas sûr que OpenAI soit du même avis !!

Thanks for your attention

More on:

AI and the
law
AI for lawyers

Alain Strowel

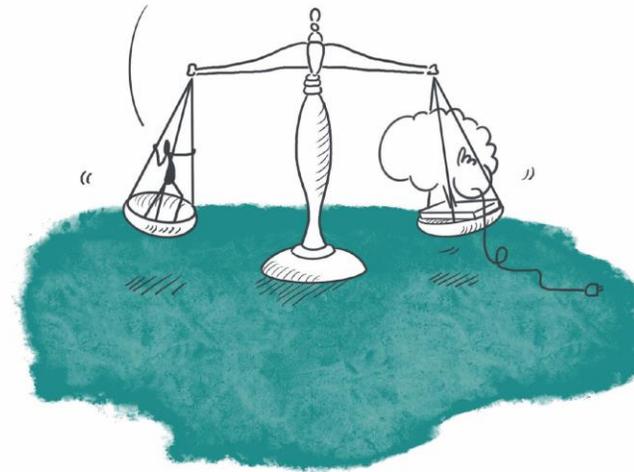
alain.strowel@uclouvain.be

Alain.strowel@pierstone.com

L'INTELLIGENCE ARTIFICIELLE POUR LES JURISTES

ALAIN STROWEL ET FRANÇOIS WÉRY

TU SAIS C'EST
JUSTE UNE QUESTION
D'ÉQUILIBRE



 LARCIER
INTERSENTIA

Q & A

QUESTIONS DE LA SALLE

LUNCH BREAK

APRAM

ECTA

Caring for a balance of IP rights

Co-organised with:



Sponsored by:



Moderator:
DELIA BELCIU

DB Law Office, RO

Speakers:

EDOUARD TREPPOZ

Panthéon - Sorbonne Paris I University, FR

ALEXANDER PEUKERT

Goethe University Frankfurt, DE

ALEXANDRA BENSAMOUN

Université Paris Saclay, FR

THOMAS BOUVET

Jones Day, FR



AI & IP: A LOVE-HATE RELATIONSHIP?

Afternoon Panel:

Data scraping to feed AI models, data protection and other IP issues

APRAM-ECTA® ETREAT JOINT EVENT (Paris, 3 April 2025, 9:00-17:00)

IA TDM : où en sommes nous?

Edouard Treppoz

Professeur Université Paris I Panthéon-Sorbonne

Training data means ...

Les « data » peuvent être des œuvres protégées par le droit d'auteur.

Faut-il alors en déduire que l'entraînement met en œuvre l'exclusivité dont bénéficie l'auteur ?

Enjeu fondamental : les titulaires de droit peuvent-ils être rémunérés lorsque leurs œuvres servent de données d'entraînement ?

Pouvons nous parler d'un acte de reproduction au sens du droit d'auteur?

Article L. 122-3 du Code de la propriété intellectuelle

« La reproduction consiste dans la fixation **matérielle** de l'œuvre par tous procédés qui **permettent de la communiquer au public de manière indirecte.** »

Pouvons nous parler d'un acte de reproduction au sens du droit d'auteur?

Oui mais en droit européen :

Article 2 Directive 2001/29 « Les États membres prévoient le droit exclusif d'autoriser ou d'interdire la reproduction directe ou indirecte, **provisoire ou permanente**, par quelque moyen et sous quelque forme que ce soit, en tout ou en partie. »

Pouvons nous parler d'un acte de reproduction au sens du droit d'auteur?

Oui mais en droit européen :

FAPL CJUE C-403/08

159 Par conséquent, il y a lieu de répondre à la question posée que l'article 2, sous a), de la directive sur le droit d'auteur doit être interprété en ce sens que le droit de reproduction s'étend aux fragments transitoires des œuvres dans la mémoire d'un décodeur satellitaire et sur un écran de télévision, à condition que ces fragments contiennent des éléments qui sont l'expression de la création intellectuelle propre aux auteurs concernés, l'ensemble composé des fragments reproduits simultanément devant être examiné afin de vérifier s'il contient de tels éléments.

Pouvons nous parler d'un acte de reproduction au sens du droit d'auteur?

Oui mais en droit voisin :

Pelham CJUE C-476/17

N° 39 Eu égard aux considérations qui précèdent, il y lieu de répondre aux première et sixième questions que l'article 2, sous c), de la directive 2001/29 doit, à la lumière de la Charte, être interprété en ce sens que le droit exclusif conféré par cette disposition au producteur de phonogrammes d'autoriser ou d'interdire la reproduction de son phonogramme lui permet de s'opposer à l'utilisation par un tiers d'un échantillon sonore, même très bref, de son phonogramme aux fins de l'inclusion de cet échantillon dans un autre phonogramme, **à moins que cet échantillon n'y soit inclus sous une forme modifiée et non reconnaissable à l'écoute.**

La réponse par l'exception?

Directive DSM 2019/790

Article 2 par 2

«fouille de textes et de données», toute technique d'analyse automatisée visant à analyser des textes et des données sous une forme numérique afin d'en dégager des informations, ce qui comprend, à titre non exhaustif, des constantes, des tendances et des corrélations;

La réponse par l'exception?

Directive DSM 2019/790

Article 3 par 1

Les États membres prévoient une exception aux droits prévus à l'article 5, point a), et à l'article 7, paragraphe 1, de la directive 96/9/CE, à l'article 2 de la directive 2001/29/CE et à l'article 15, paragraphe 1, de la présente directive **pour les reproductions et les extractions effectuées par des organismes de recherche et des institutions du patrimoine culturel, en vue de procéder, à des fins de recherche scientifique, à une fouille de textes et de données sur des œuvres ou autres objets protégés auxquels ils ont accès de manière licite.**

La réponse par l'exception?

Directive DSM 2019/790

Article 4

1. Les États membres prévoient une exception ou une limitation (...) pour les reproductions et les extractions d'œuvres et d'autres objets protégés accessibles de manière licite aux fins de la fouille de textes et de données. (...)

3. L'exception ou la limitation prévue au paragraphe 1 s'applique à condition que l'utilisation des œuvres et autres objets protégés visés audit paragraphe **n'ait pas été expressément réservée par leurs titulaires de droits** de manière appropriée, notamment par des procédés lisibles par machine pour les contenus mis à la disposition du public en ligne.

La réponse par l'exception ?

Oui mais ...

L'article 4 est-il applicable à des IAG non connues du législateur en 2019?

La définition de l'exception « afin d'en dégager des informations, ce qui comprend, à titre non exhaustif, des constantes, des tendances et des corrélations » correspond-elle aux IAG?

L'interprétation authentique *a posteriori* par le règlement IA (2024/1689)

Article 53

1) Les fournisseurs d'IA à usage général :

c) mettent en place une politique visant à se conformer au droit de l'Union en matière de droit d'auteur et droits voisins, et notamment à identifier et à respecter, y compris au moyen de technologies de pointe, une réservation de droits exprimée conformément à l'article 4, paragraphe 3, de la directive (UE) 2019/790;

L'interprétation authentique *a posteriori* par le règlement IA

Considérant 105

La directive (UE) 2019/790 a introduit des exceptions et des limitations autorisant les reproductions et extractions d'œuvres ou d'autres objets protégés aux fins de la fouille de textes et de données, sous certaines conditions. En vertu de ces règles, les titulaires de droits peuvent choisir de réserver leurs droits sur leurs œuvres ou autres objets protégés afin d'empêcher la fouille de textes et de données, à moins que celle-ci ne soit effectuée à des fins de recherche scientifique. **Lorsque les droits d'exclusion ont été expressément réservés de manière appropriée, les fournisseurs de modèles d'IA à usage général doivent obtenir une autorisation des titulaires de droits s'ils souhaitent procéder à une fouille de textes et de données sur ces œuvres.**

Les difficultés théoriques : une formalité prohibée?

L'opt-out est-il conforme à la Convention de Berne prohibant toute formalité au titre de la jouissance et de l'exercice des droits?

Pour rappel : CJUE, 16 novembre 2016, C-301/15 (à propos de la loi FR Relire) 50 il résulte de l'article 5, paragraphe 2, de la convention de Berne, qui s'impose à l'Union pour les raisons exposées au point 32 du présent arrêt, que la jouissance et l'exercice des droits de reproduction et de communication au public attribués aux auteurs par cette convention et correspondant à ceux prévus à l'article 2, sous a), et à l'article 3, paragraphe 1, de la directive 2001/29 ne peuvent être assujettis à aucune formalité.

Les difficultés théoriques : une exception contraire au test en trois étapes?

Article 5 Directive 2001/29

5. Les exceptions et limitations prévues aux paragraphes 1, 2, 3 et 4 ne sont applicables que dans certains cas spéciaux qui ne portent pas atteinte à l'exploitation normale de l'œuvre ou autre objet protégé ni ne causent un préjudice injustifié aux intérêts légitimes du titulaire du droit.

CJUE, 16 juin 2011, Stichting de Thuiskopie, C-462/09

22 Ainsi, eu égard à cette dernière condition, les États membres, lorsqu'ils décident d'instaurer l'exception de copie privée dans leur droit national, sont, en particulier, tenus de prévoir, en application de l'article 5, paragraphe 2, sous b), de la directive 2001/29, le versement d'une «compensation équitable» au bénéfice des titulaires des droits

Quid de l'opt-out sur la mise en œuvre du triple test?

Les difficultés théoriques : vers des paradis pour IAG?

Vers un risque de délocalisation de l'entraînement :

Considérant 106 règlement IA

Tout fournisseur qui met un modèle d'IA à usage général sur le marché de l'Union devrait se conformer à cette obligation, quelle que soit la juridiction dans laquelle se déroulent les actes pertinents au titre du droit d'auteur qui sous-tendent l'entraînement de ces modèles d'IA à usage général. Cela est nécessaire pour garantir des conditions de concurrence équitables entre les fournisseurs de modèles d'IA à usage général, lorsqu'aucun fournisseur ne devrait pouvoir obtenir un avantage concurrentiel sur le marché de l'Union en appliquant des normes en matière de droit d'auteur moins élevées que celles prévues dans l'Union.

Les difficultés pratiques : qui ?

Article 4 DSM

n'ait pas été expressément réservée par **leurs titulaires de droits ...**

Article 122-5-3 du Code de la Propriété intellectuelle

III.-Sans préjudice des dispositions du II, des copies ou reproductions numériques d'œuvres auxquelles il a été accédé de manière licite peuvent être réalisées en vue de fouilles de textes et de données menées à bien par toute personne, quelle que soit la finalité de la fouille, **sauf si l'auteur s'y est opposé de manière appropriée, notamment par des procédés lisibles par machine pour les contenus mis à la disposition du public en ligne.**

Les difficultés pratiques : comment ?

Article 4 DSM

n'ait pas été expressément réservée par leurs titulaires de droits **de manière appropriée, notamment par des procédés lisibles par machine** pour les contenus mis à la disposition du public en ligne.

Des premières interprétations divergentes:

Hambourg Landgericht, 27 septembre 2024, LAION v Rechbank
Amsterdam 30 octobre 2024, DPG Media & Mediahuis / HOWARDSHOME

La difficulté diabolique : la preuve impossible?

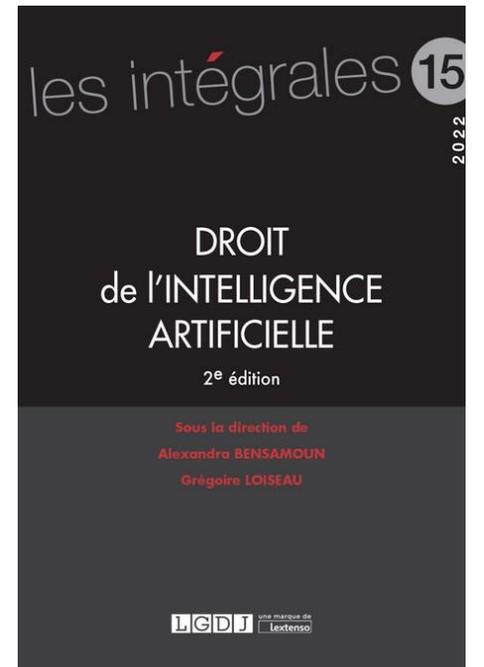
- Obligation de transparence et preuve?
- Vers des aménagements probatoires nationaux?

AI & IP: a love-hate relationship?
APRAM-ECTA, 3 avril 2025

université
PARIS-SACLAY

Transparence et preuve de l'utilisation d'un contenu protégé par l'IA

Alexandra BENSAMOUN
*Professeure de droit, Université Paris-Saclay
Personnalité qualifiée au CSPLA (min. Culture)*



A LA MANIÈRE DE...



Une question de souveraineté(s)

Vous avez dit IA ? Pas seulement de la tech' !

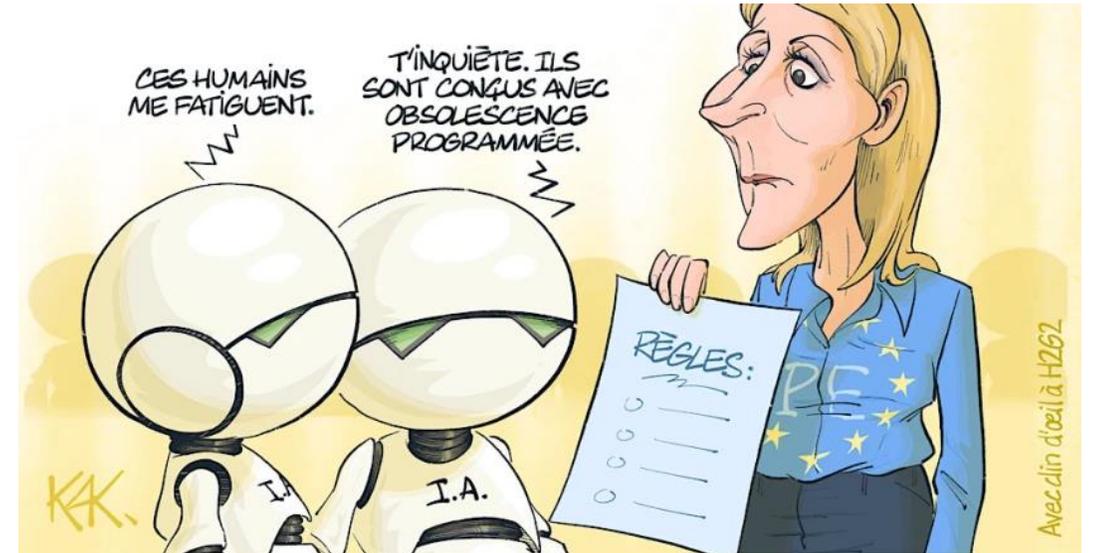
Souveraineté normative

Souveraineté économique

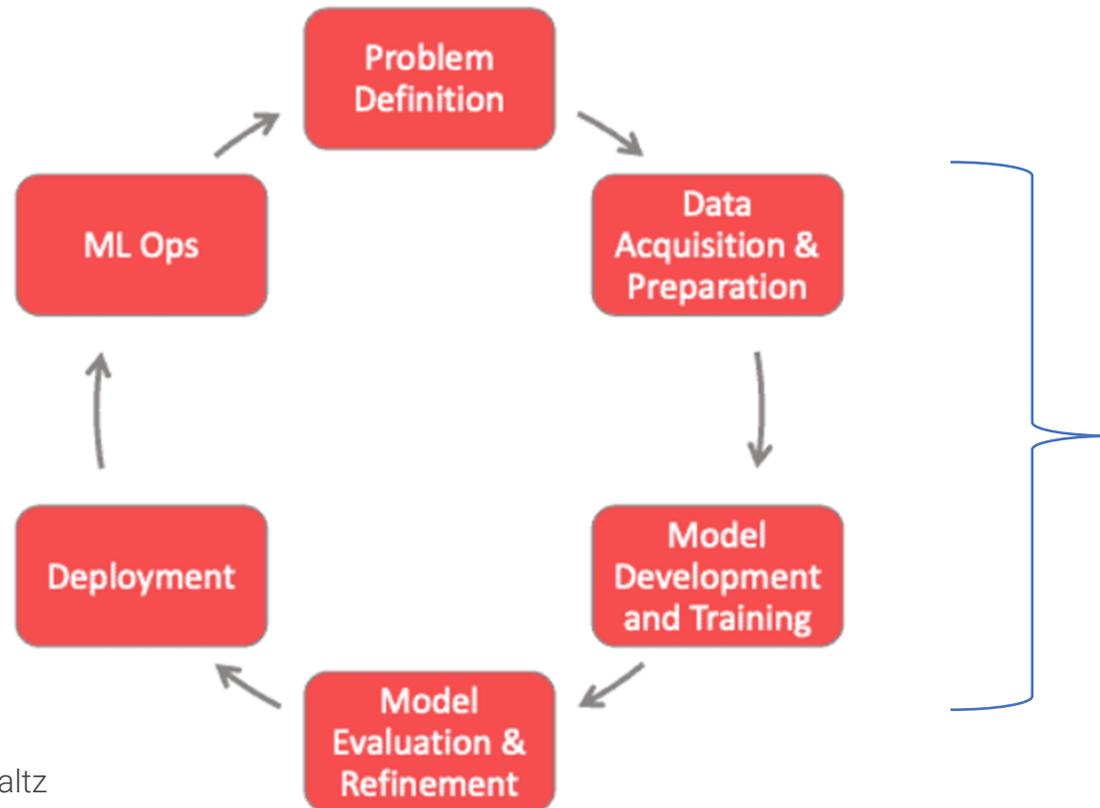
Souveraineté culturelle

L'opposition stérile entre innovation et régulation

La nécessité d'un cadre de confiance, qui respecte les droits et la chaîne de valeur – un **modèle européen**



GenAI



Web scraping - Web crawling

GenAI Lifecycle, Jeff Saltz



Comment rendre au droit son effectivité ?

Opacité des opérations : quels contenus utilisés, pour quels usages, conditions de l'exception respectées (accès licite et *opt out*) ?

Seul le fournisseur d'IA peut « ouvrir le capot ».

AI Act : obligation de transparence



Art. 53, 1, c et d

« Les fournisseurs de modèles d'IA à usage général : (...)

c) mettent en place une **politique visant à se conformer** au droit de l'Union en matière de droit d'auteur et droits voisins, et notamment à identifier et à respecter, y compris au moyen de technologies de pointe, une réservation de droits exprimée conformément à l'article 4, paragraphe 3, de la directive (UE) 2019/790 ;

d) élaborent et mettent à la disposition du public un **résumé suffisamment détaillé** du contenu utilisé pour entraîner le modèle d'IA à usage général, conformément à un modèle fourni par le Bureau de l'IA. »



Les deux obligations

- Mise en place d'une « *compliance by design* » (pt c)
Mise en œuvre de mesures permettant le respect de l'acquis communautaire en DA/DV ; « technologies de pointe »
+ *Code of practice* en cours : V3 critiquée par les titulaires de droits (prise de position), « efforts raisonnables »
- Elaboration et mise à la disposition du public d'un « résumé suffisamment détaillé » des contenus utilisés pour l'entraînement des modèles (pt d) : oxymore, lobbying intense
+ « Template » fourni par le Bureau de l'IA

Entrée en vigueur en août 2025

Joint statement by a coalition of authors, performers and other rightsholders active across the EU's cultural and creative sectors regarding the third draft of the EU AI Act's GPAI Code of Practice

The third draft of the GPAI Code of Practice undermines the objectives of the AI Act, contravenes EU law and ignores the intention of the EU legislator – we cannot support it.



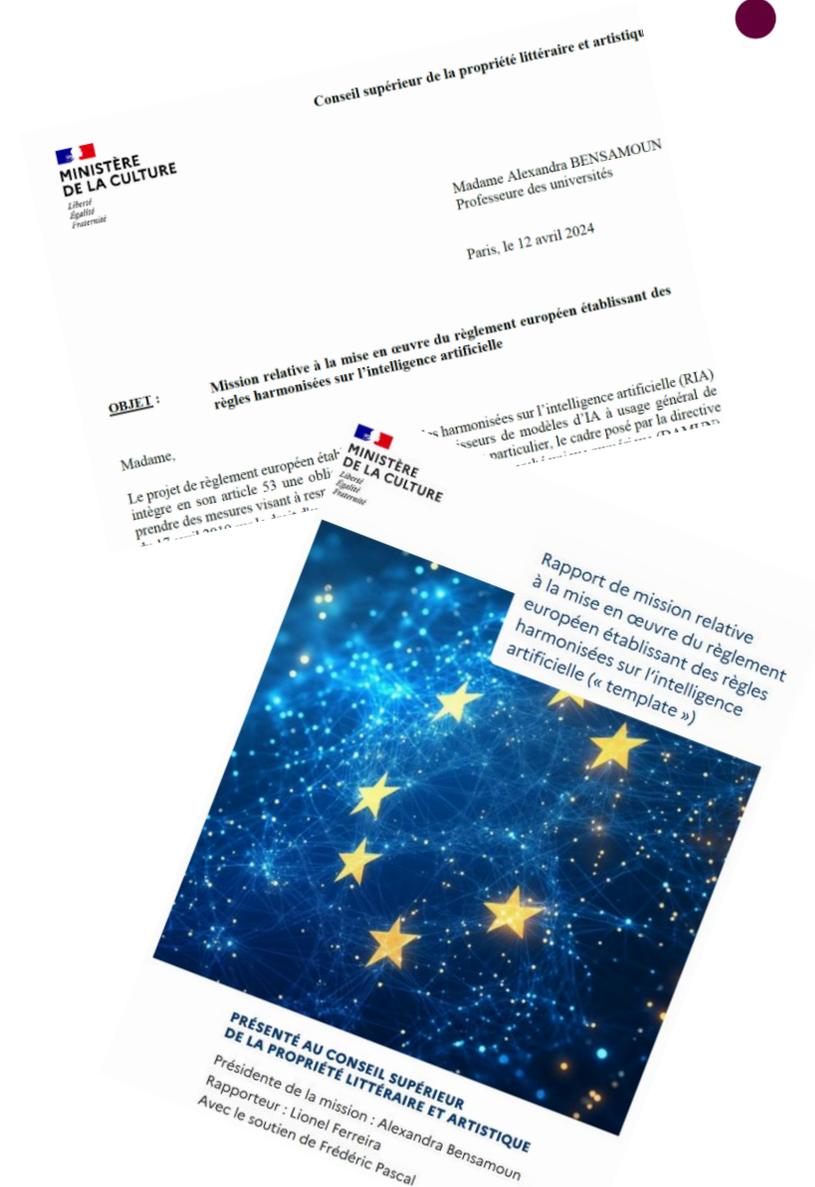
The intention of the EU's AI Act was to enable the development of responsible AI and ensure that the cultural and creative sectors can continue to grow across European territories. Entirely counter to this, the third draft of the GPAI Code of Practice sets an implementation standard which undermines the objectives of the AI Act, contravenes EU law and ignores the intention of the EU legislator.

The third draft is therefore completely unacceptable. We cannot support the GPAI Code of Practice in its current version. No Code would be better than the fundamentally flawed third draft.



Interprétation ?

1. Lien entre les deux obligations : l'objectif de transparence
2. Interprétation contextuelle :
 - Une disposition finalisée : l'exercice et le respect des droits : « *aider les parties ayant des intérêts légitimes, y compris les titulaires de droits d'auteur, à exercer et à faire respecter les droits que leur confère la législation de l'Union* »
 - Une disposition contrainte : le respect du secret des affaires (pb de la mise à disposition du public)
3. « *Effet utile* » // CJUE





This is a template. Make a copy before proceeding.

Gemini Training Data for [nickname] Model

Attorney Client Privileged

Once you create your data card doc, PLEASE IMMEDIATELY SHARE VIEWER ACCESS TO [redacted]@google.com, and then delete this box.

Authors & Contributors: [redacted]@google.com

Created: 20 Feb 2024 (update date)
 Status: Draft
 Self link: (add go/ins)
 Updated: 6 Sept 2023 (update date)

Context

This document describes the data used for training the [nickname] model and tracks LGTMs from project leads and data owners based on known use cases as of the creation date. Any future product use cases should go through the appropriate data & model governance process. See Use Cases section

LGTMs (based on details in Training Data section and approvals already granted via the GDM Legal Review process)

Reviewer	Reviewer Role	LGTM	LGTM Date	Comments/Notes
Tom Lue	General Counsel, Google DeepMind			
Rich Lee	Product Counsel, Search			
Jeff Dean Koray Kavukcuoglu Orio Vinyals	Gemini Leadership			
Megha Goel	Gemini PM			
	[add reviewers for additional data owners, as required]			
	[add reviewers for additional project leads, as required]			

Data sources	Data contained from the data source	Included in Gemini v1 IT model(s) (and any changes)	Prior approvals	Filtering	Contains PHI?	Contains User Data?	Contains licensed data?	Contains Google proprietary data?	Citation attribution in the Dataset (e.g. URL, License, etc)?	Planned mitigations
Text										

Training Data

Summary of Training Data
 [Add a summary of training data, including links to data cards for models contained within this new model]

#	Dataset Name + link to MLC entry	Data contained in the dataset (description)	Link to approval	Link to any prior approvals / precedence with this dataset	Paths to datastore that holds the train + test sets used for the model (separate multiple values with a comma)	Contains PII? (if so, provide detail)	Contains User Data?	Contains Public Data?	Subject to retention? If so, link to PDD + wipeout plan.	Contains data from a 3P via contract (e.g. Shutterstock images)?	Contains Open Source data?	Citation attribution / license	Did we filter or alter the dataset for Gemini?	Filtering process (and any other data pre-processing), in as much detail as possible. Also highlight any deviations from normal process (e.g. no link filtering, etc)	Required Mitigations	Risks

Use Cases

[edit this section based on use case limitations / particulars of this model]

Approval of the data in this card means that Gemini models trained on this data can be used widely across Google PAs including Cloud, for research and productization.

- <checkpoint name> is approved for
 - Example: Product launches or previews that have been approved by Gemini Leads, including, but not limited to, Cloud releases
 - Example: Internal testing
- <checkpoint name> is not approved for
 - Example: Product launches or previews that have not been approved by Gemini Leads

Interprétation ?

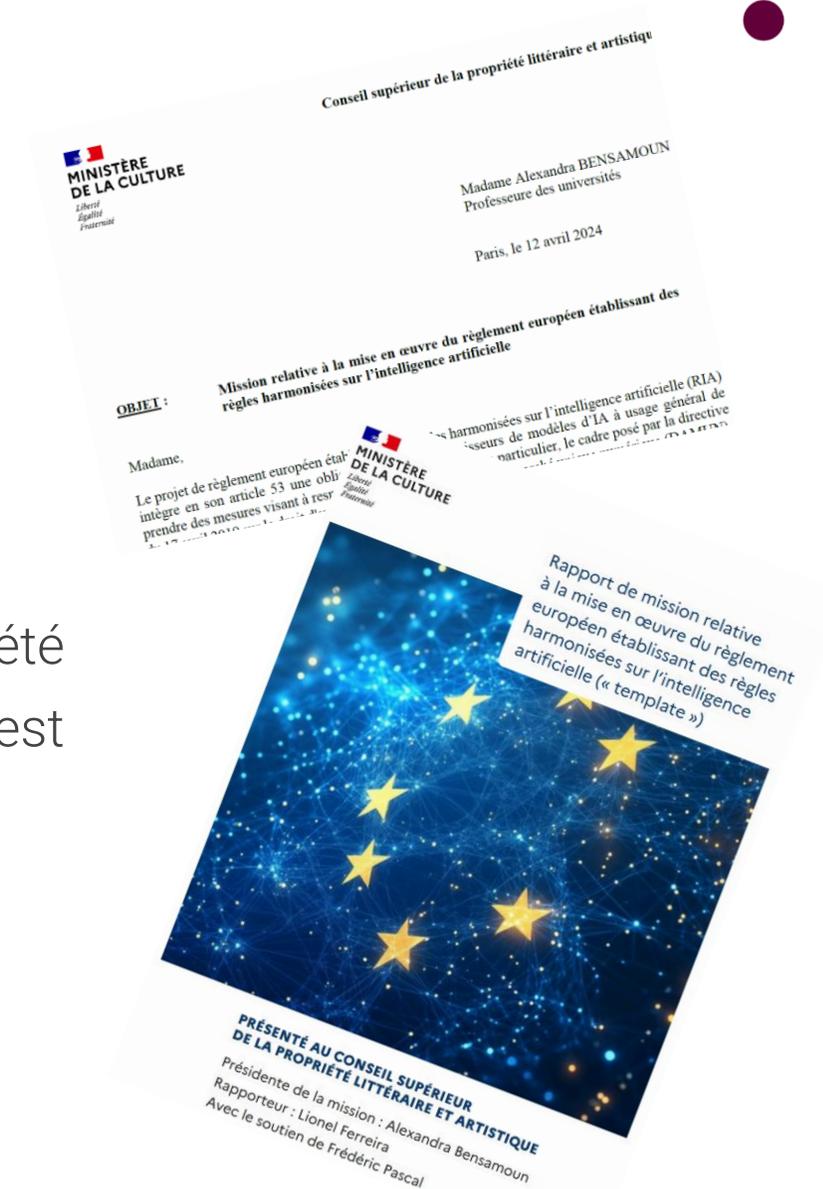
= divulguer « les ingrédients mais pas la recette » (Rapport CSPLA)
 (« *complet en termes de contenu* »)

Sauf que certains contenus pourraient finalement ne pas avoir été utilisés et le dire « au public » (et donc aux concurrents) est problématique, tout comme ne pas le dire aux titulaires de droits !

Etape d'après ???

Négociation directe ? Dialogue ?

Autorité ?



Transparence ?

Transparence au-delà du DA et des DV (autres données, autres dispositions et autres textes)

= un « principe matriciel »

L'opacité empêche l'application du droit par le juge et le contrôle de conformité du régulateur.

Sans transparence, il est impossible de vérifier que le droit est respecté (notamment les conditions de l'exception).



« le respect d'un droit en général (ou la preuve de son atteinte) en cas d'usage de systèmes d'IA ne saurait être garanti sans transparence au risque de se muer en simple déclaration de principe. Le lien entre l'effectivité et la transparence est alors essentiel. **Et plus fondamentalement, sans transparence, c'est au droit à un recours effectif que l'on attende** ».

Transparence et marché

La transparence est le préalable à l'émergence d'un **marché éthique et compétitif**, qui respecte la chaîne de valeur.

Constat : la rencontre des besoins

Note d'étape juridique, déc. 2024 : rappel des principes (exclusivité, liberté contractuelle), quelles incitations (faciliter l'accès aux données ?), quelle valeur (critères)... ?

Rapport, fin avril-début mai 2025



OBJET : Mission relative à la rémunération des contenus culturels utilisés par les systèmes d'intelligence artificielle

Mesdames,

Les technologies d'apprentissage automatique qui sont au cœur des systèmes d'intelligence artificielle (IA) reposent sur d'importantes masses de données culturelles.

S'agissant des données protégées par des droits de propriété littéraire et artistique, les titulaires de droits français, confrontés à l'utilisation croissante de leurs œuvres et prestations, décident de plus en



Transparence et droits fondamentaux

L'absence de transparence attente aux droits fondamentaux, précisément au droit au recours et au droit de propriété.

Com., 5 février 2025 : « le droit à la preuve peut justifier la production d'éléments couverts par le secret des affaires ».

Quelles solutions ? Comment recouvrer la capacité à apporter la preuve ? Mécanismes nationaux de facilitation probatoire ?

Droit commun – saisie-contrefaçon, droit d'information, 145 CPC... Réforme ? Proposition de présomption



« le respect d'un droit en général (ou la preuve de son atteinte) en cas d'usage de systèmes d'IA ne saurait être garanti sans transparence au risque de se muer en simple déclaration de principe. Le lien entre l'effectivité et la transparence est alors essentiel. Et plus fondamentalement, sans transparence, c'est au droit à un recours effectif que l'on attende ».

Are U a JedAI ?

« L'homme est lent, peu rigoureux et très intuitif.

L'ordinateur est super rapide, très rigoureux et complètement con ».

Gérard Berry
Professeur au Collège de France
Titulaire de la chaire Algorithmes, machines et langages



IP ASPECTS LINKED TO ARTICLE 53 AI ACT: TRANSPARENCY OBLIGATION AND HOW TO PROVE THAT AI PROVIDERS HAVE USED YOUR WORK

3 avril 2025

Thomas Bouvet



JONES
DAY®

SOMMAIRE

1. Introduction: Place des secrets d'affaires en matière d'IA
2. Obligations de transparence du Règlement sur l'IA
3. Comment prouver l'utilisation d'œuvres protégées ?

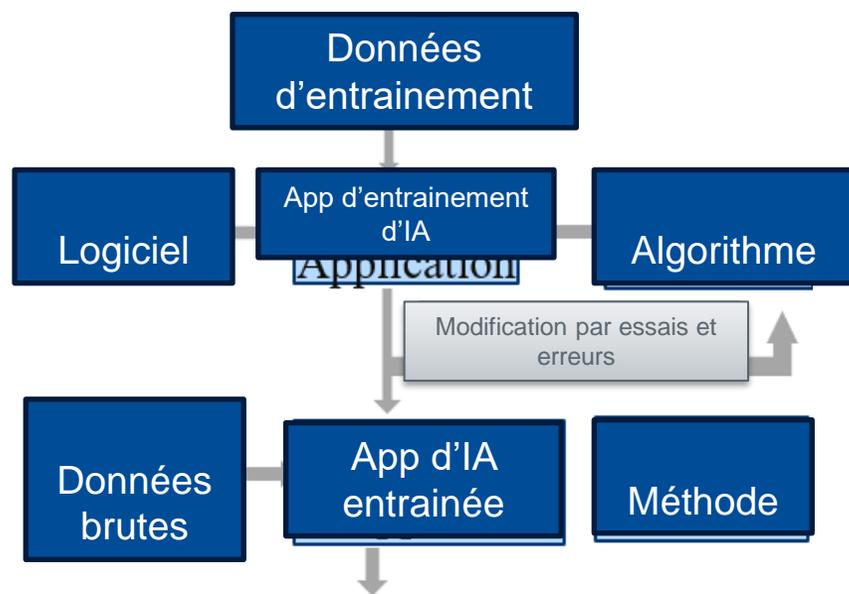
PLACE DE LA PROTECTION PAR SECRETS D'AFFAIRES EN MATIÈRE D'OUTILS D'IA

CONDITIONS DE PROTECTION PAR SECRETS D'AFFAIRES SELON LA DIRECTIVE SUR LES SECRETS D'AFFAIRES ET LE DTSA (US)



- **Secret** : L'information n'est pas généralement connue ou facilement accessible par d'autres (DTSA US : ou facilement accessible par des moyens légitimes)
- **Valeur commerciale** : L'information a une valeur commerciale du fait de son caractère secret
- **Mesures / efforts raisonnables** : L'information a fait l'objet de mesures raisonnables pour maintenir sa confidentialité

PLACE DES SECRETS D'AFFAIRES ET DES DROITS DE PROPRIÉTÉ INTELLECTUELLE EN MATIÈRE D'IA



	Secrets d'affaires	Brevets	Droit d'auteur
Données d'entraînement, de test, de validation	✓		Droit d'auteur, bases de données, logiciels, etc.
Logiciels et algorithmes	✓	(✓)	✓
Méthodes et techniques d'entraînement	✓	✓	
Produits et services d'IA	✓	✓	✓
Résultats	✓	✓	✓

JURISPRUDENCE : PROTECTION DE DONNÉES D'ENTRAÎNEMENT

- Tribunal de Milan, 14 mai 2018, *Leonardo Assicurazioni s.r.l. c. Pro Insurance s.r.l. & A&A Insurance Broker*:
*"Leonardo Assicurazioni est un agent général de la compagnie d'assurance Generali Italia à Milan. Il a acquis **un patrimoine d'informations secrètes et sensibles qui sont maintenant stockées et protégées par un logiciel appelé SIAG**. Grâce à un système intelligent, le logiciel SIAG sélectionne automatiquement les produits en fonction des besoins des clients, dans une exploitation économique de la base de données. (...) la probabilité de succès sur le fond a été confirmée par un nombre pertinent de dossiers concernant les portefeuilles de clients qui correspondaient aux produits associés, tels que traités par la technologie IA du logiciel, et que **cette combinaison particulière de données conférait un avantage commercial massif au propriétaire du secret commercial**. (...) la valeur et l'accessibilité de ces informations doivent être évaluées non pas par rapport aux données individuelles isolées, mais à la configuration et à la combinaison des éléments de celles-ci."*
- Rapport 2023 de l'EUIPO sur les secrets d'affaires concernant cette décision : "Ce cas doit être considéré comme un exemple de la manière dont l'acquisition illégale de secrets d'affaires peut affecter l'économie des données, où l'information secrète est une source de valeur commerciale massive non pas en elle-même, mais en raison de la manière dont elle peut être traitée par les technologies IA."

JURISPRUDENCE : PROTECTION DE BASES DE DONNÉES

- CA Montpellier, 14 mai 2019, *Vestergaard v. IIC, O. Skovmand*, RG n° 15/07646:
une base de données contenant les résultats de plus de cinq ans de R&D (y compris le **savoir-faire négatif** : essais et erreurs) dans le domaine des moustiquaires en polyéthylène imprégnées d'insecticide à longue durée.
- *Genentech, Inc. c. JHL Biotech, Inc.*, No. C 18-06582 WHA, 2019 WL 1045911 (N.D. Cal. Mar. 5, 2019):
*« Les secrets d'affaires en question concernent **des informations propriétaires non divulguées concernant la manière dont Genentech a décidé de la formulation la plus appropriée pour [le médicament biologique]**, et les forces et faiblesses relatives des diverses formulations pour [le biologique] que Genentech a envisagées. Ce type d'informations propriétaires peut prendre la forme de **savoir-faire négatif**, et la propre formulation de JHL peut encore avoir bénéficié de ces informations en évitant des écueils de recherche coûteux. »*

JURISPRUDENCE : PROTECTION D'OUTILS D'IA

- Communiqué de presse du ministère de la Justice des États-Unis, un ressortissant chinois résidant en Californie arrêté pour vol de secrets commerciaux liés à l'intelligence artificielle de Google (*USA c. Linwei Ding*, No. 3:24-cr-00141 VC (N.D. Cal. 5 mars 2024):

« Le procureur général Garland a déclaré "Dans ce cas, nous alléguons que le défendeur a volé des secrets commerciaux liés à l'intelligence artificielle de Google tout en travaillant secrètement pour deux entreprises basées en Chine. Nous protégerons féroce­ment les technologies sensibles développées en Amérique contre ceux qui ne devraient pas les avoir. (...) »

- CA Paris, 20 mai 2021, *Nobel Connexion c. The Call Machine*, RG n° 20/04388

Logiciel permettant de réaliser des calculs complexes et des prédictions sur l'applicabilité des avantages légaux, même si ceux-ci sont fournis conformément aux dispositions légales connues.

L'ENJEU DES DONNÉES D'ENTRAÎNEMENT

- « Jeux de données », se réfère aux « Données d'entraînement », « Données de test » et « Données de validation », à savoir à toute information utilisée dans le cadre du développement et de l'exploitation de tout logiciel basé sur l'IA, plus particulièrement à toute donnée utilisée pour « entraîner » un modèle d'IA générative
- L'entraînement des modèles d'IA est une étape critique du développement d'un outil d'IA car il influence directement la qualité globale du modèle, représente une part significative des coûts totaux associés au développement et à la maintenance des modèles d'IA, et dans le contexte des données ouvertes et des ressources partagées, la sélection privée d'un ensemble de données brutes disponibles publiquement constitue un facteur de différenciation clé face à la concurrence
- La qualité d'un produit IA dépend d'une combinaison de la qualité des ensembles de données utilisés par un fournisseur d'IA, y compris la sélection et la curation de l'ensemble de données, la quantité de données dans les ensembles de données ainsi que la conception de l'outil IA lui-même
- La divulgation non réglementée et/ou excessive des ensembles de données pourrait compromettre injustement la compétitivité de leurs détenteurs/conservateurs et annihiler les investissements réalisés par eux

OBLIGATION DE TRANSPARENCE PRÉVUE PAR LE RÈGLEMENT SUR L'IA

RÈGLEMENT (EU) SUR L'IA DU 13 JUIN 2024

- Le Règlement sur l'IA impose une série d'obligations dont l'application dépend :
 - (i) de la qualification du risque soulevé par les systèmes IA (risque inacceptable, risque élevé, risque systémique, risque limité et risque minimal ou nul) et
 - (ii) du niveau dans la chaîne de valeur (des fournisseurs aux déployeurs de systèmes IA).
- Le non-respect peut entraîner des sanctions, y compris des ordres ou des avertissements, et des amendes administratives
- La plupart des dispositions sont applicables à partir du 2 août 2026
- Le Règlement sur l'IA s'applique aux entités situées à l'intérieur et à l'extérieur de l'UE, dès lors que (i) le système d'IA est mis sur le marché de l'UE, (ii) le déployeur a son lieu d'établissement ou est situé dans l'UE, ou (iii) la sortie produite par le système d'IA est utilisée dans l'UE

OBLIGATIONS DE DIVULGATION IMPOSÉES AUX FOURNISSEURS DE GPAI

Pour l'office de l'IA et les offices nationaux

Art 53(1)(a) RIA :

Documentation technique du modèle, y compris son processus d'entraînement et d'essai (y compris des informations sur les données utilisées pour l'entraînement) et les résultats de son évaluation

Pour les fournisseurs aval

Art 53(1)(b) RIA :

Informations permettant aux fournisseurs de systèmes IA, intégrant GPAI, d'avoir une bonne compréhension des capacités et des limites du modèle et de se conformer à leurs obligations et, au minimum, à l'annexe XII

Pour respecter le droit d'auteur

Art 53(1)(c) RIA :

Une politique pour se conformer à la législation de l'Union sur le droit d'auteur et les droits connexes, et en particulier pour identifier et se conformer aux opt-outs de l'exception de fouille de données.

Pour le public

Art 53(1)(d) EU RIA :

Un résumé suffisamment détaillé sur le contenu utilisé pour l'entraînement du modèle d'IA à usage général, selon un modèle fourni par l'Office de l'IA.



Tout en tenant dûment compte de la nécessité de protéger les secrets d'affaires et les informations commerciales confidentielles, ce résumé devrait être généralement complet en termes de contenu plutôt que détaillé sur le plan technique afin d'aider les parties ayant des intérêts légitimes, y compris les titulaires de droits d'auteur, à exercer et à faire respecter les droits que leur confère la législation de l'Union, par exemple en énumérant les principaux jeux ou collections de données utilisés pour entraîner le modèle, tels que les archives de données ou bases de données publiques ou privées de grande ampleur, et en fournissant un texte explicatif sur les autres sources de données utilisées.

(Considérant 107 EU AI Act)



ÉQUILIBRE DES RAISONS D'ÊTRE

- L'Accord ADPIC et la Directive sur les secrets d'affaires prévoient que la limitation des secrets d'affaires doit être aussi encadrée que possible (Art. 7 ADPIC)
- Les obligations de transparence du Règlement IA sont soumises à la protection des secrets d'affaires (Art. 53(1)-(7) et 78 Règlement IA) :
 - « *Tout en tenant dûment compte de la nécessité de protéger les secrets d'affaires et les informations commerciales confidentielles* »,
 - « *Sans préjudice de la nécessité de respecter et de protéger les droits de propriété intellectuelle et les informations de nature commerciale ou les secrets d'affaires* »,
 - « *Les autorités (...) ne doivent demander que les données strictement nécessaires...* ».
- La divulgation des données d'entraînement a pour but « *d'aider les parties ayant des intérêts légitimes, y compris les titulaires de droits d'auteur, à exercer et à faire respecter les droits que leur confère la législation de l'Union* » (Art. 53(1)(d) IA Acte)

TEST D'ÉQUILIBRE DES TEXTES

- Le test d'équilibre de la CJUE implique une analyse en trois étapes
 - la mesure est-elle adaptée à la réalisation d'un objectif légitime (« critère de légitimité »)?
 - existe-t-il des moyens moins restrictifs disponibles (« critère de nécessité »)?
 - la mesure n'impacte-t-elle pas excessivement les intérêts légitimes (« critère de proportionnalité »)?
- Affaires jointes C-453/03, C-11/04, C-12/04 et C-194/04 ABNA Ltd et al. c. Secretary of State for Health *et al.*:
 - l'obligation de divulguer la composition précise des aliments pour animaux, imposée par la législation communautaire pour protéger contre la contamination du cycle alimentaire, porte inutilement atteinte aux intérêts économiques des fabricants, obligeant ces derniers « à divulguer les formules de composition de leurs produits, au risque que ces produits soient utilisés comme modèles, éventuellement par ces clients eux-mêmes » et parce que « les fabricants ne peuvent pas tirer profit des investissements qu'ils ont réalisés en termes de recherche et d'innovation » ;
 - la CJUE suggère que l'indication d'une fourchette pour les composants suffit à des fins d'information.

APPLICATION DU TEST D'ÉQUILIBRE

- Art 53(1)(c) RIA : le résumé a pour but « ***d'aider les parties ayant des intérêts légitimes, y compris les titulaires de droits d'auteur, à exercer et à faire respecter les droits que leur confère la législation de l'Union*** »:
 - « aider »: le résumé ne doit pas « *permettre* » aux titulaires de droits d'auteur de faire valoir leurs droits ; il doit être compris comme fournissant uniquement des informations générales permettant aux titulaires de droits d'auteur de savoir s'ils doivent demander des précisions
 - seules les « parties ayant des intérêts légitimes » sont habilitées à demander et à obtenir les informations nécessaires à l'exercice de leurs droits, ce qui exclut le grand public
- Il n'est ni légitime ni nécessaire de contraindre les fournisseurs de modèles d'IA à divulguer au grand public des informations détaillées sur les ensembles de données qu'ils utilisent pour entraîner leurs modèles d'IA et sur leur contenu

CODE DE BONNES PRATIQUES ET OBLIGATIONS DE TRANSPARENCE : DE QUOI IL S'AGIT ?

Qu'est-ce que le code de bonnes pratiques pour les modèles d'IA à usage général ?

- Le code de bonnes pratiques est un **outil volontaire** pour aider les fournisseurs de GPAI à se conformer au RIA. Il repose sur des engagements et de mesures portant, entre autres, sur la **transparence** et le respect du **droit d'auteur**.
- Une **troisième version** du document a été élaboré par des experts indépendants sous la supervision du bureau de l'IA de la Commission européenne:
 - inclut un **formulaire de documentation** que les fournisseurs de modèles devront remplir
 - le respect des engagements en matière de droit d'auteur doit être proportionné à la taille et aux capacités des fournisseurs
 - les fournisseurs sont encouragés –non plus obligés – à publier un résumé de leur politique de respect du droit d'auteur
- Une quatrième et dernière version est attendue d'ici le 2 mai 2025
 - Vers une volonté politique plus générale de déréglementation (gouvernement allemand et initiative « D9 + »)

CODE DE BONNES PRATIQUES ET OBLIGATIONS DE TRANSPARENCE : ÉTAT ACTUEL DES OBLIGATIONS DE TRANSPARENCE ?

Quelles sont les obligations de transparence dans le code de bonnes pratiques ?

- **Rédiger et mettre à jour** une documentation détaillée incluant les ressources informatiques et la consommation d'énergie, conservée pendant **10 ans après la mise sur le marché**
 - proportionnées à la taille et aux capacités des fournisseurs
 - Exemption des modèles open source (sauf si risque systémique)
- **Fournir aux autorités et aux fournisseurs en aval** les informations nécessaires pour évaluer la conformité des modèles et permettre leur intégration dans des systèmes d'IA.
- **Garantir la qualité, l'intégrité et la sécurité** des informations documentées en les protégeant contre les modifications involontaires et en les conservant comme preuve de conformité.
- **Encourager la transparence**, notamment en rendant certaines données accessibles au public pour favoriser une meilleure compréhension des modèles d'IA
 - limités aux principaux *datasets* dépassant un seuil (quid accords de licence sur données)
 - Quid des URLs / noms de domaines fouillés et la fréquence des fouilles en ligne?
 - Quid des procédés destinés à retirer les données protégées?

MODÈLE POUR LE RÉSUMÉ DES DONNÉES DE FORMATION (17 JANVIER 2025)

1.1 Identification du modèle et du fournisseur

- Nom et coordonnées du prestataire
- Identifiant du modèle
- Représentant autorisé
- Modèle(s) de base

1.2. Date de mise sur le marché et date limite de connaissance

1.3. Taille, modalités et caractéristiques globales des données de formation

Modalités	Taille totale
<input type="checkbox"/> Texte	Nombre de jetons ou d'octets
<input type="checkbox"/> Image	Images numériques (ou paires avec autres médias)
<input type="checkbox"/> Vidéo	Nombre de minutes (ou de paires avec autres médias)
<input type="checkbox"/> Audio	Nombre de minutes (ou de paires avec autres médias)
<input type="checkbox"/> Autres	_____ [veuillez préciser]

Description des caractéristiques linguistiques, régionales, démographiques et autres caractéristiques pertinentes de l'ensemble des données de formation :

Texte	Image	Vidéo	Audio
<input type="checkbox"/> Textes de fiction, littérature <input type="checkbox"/> Textes scientifiques et éducatifs <input type="checkbox"/> Actualités, journalisme et avis <input type="checkbox"/> Documents juridiques et officiels <input type="checkbox"/> Communication sociale (par exemple, messages) <input type="checkbox"/> Promotion, publicité, l'examen des produits et des services <input type="checkbox"/> Autres textes	<input type="checkbox"/> Photographie <input type="checkbox"/> Peintures et beaux-arts <input type="checkbox"/> Infographie <input type="checkbox"/> Illustration et graphisme conception <input type="checkbox"/> Images sociales / personnelles Spécial <input type="checkbox"/> Code source <input type="checkbox"/> Données structurées (par exemple, calendrier, cartes)	<input type="checkbox"/> Films, spectacles, représentations <input type="checkbox"/> Contenu vidéo animé <input type="checkbox"/> Jeux vidéo et séquences immersives (par exemple, 3D) <input type="checkbox"/> Documentaires <input type="checkbox"/> Actualités et journalisme vidéo <input type="checkbox"/> Contenu utilisateur, courtes vidéos <input type="checkbox"/> Autres contenus vidéo (par exemple, art expérimental, effets vidéo)	<input type="checkbox"/> Musique <input type="checkbox"/> Narration et fiction (par exemple, livres audio) <input type="checkbox"/> Contenu audio éducatif non fictionnel <input type="checkbox"/> Émissions de radio et podcasts <input type="checkbox"/> Communication sociale (appels téléphoniques, messages vocaux) <input type="checkbox"/> Autres (par exemple, sons et ambiance)



COMMENT PROUVER L'UTILISATION D'ŒUVRES PROTÉGÉES ?

DES MODES DE PREUVE DE DROIT COMMUN

- Mesures de droit commun :
 - droit d'information (article L 331-1-2 du CPI)
 - mesures d'instruction (article 145 du CPC)
 - saisie-contrefaçon (article L. 332-1 du CPI): recherche de métadonnées, traces numériques et journaux / historiques d'utilisation

DES MODES DE PREUVE DÉDIÉS: UTILISATION D'OUTILS DE *SCRAPPING* / *REVERSE ENGINEERING*

- Différents outils permettent de rechercher si une œuvre originale est susceptible d'avoir été utilisée pour entraîner un outil d'IA:
 - Utilisation du site Have I been Trained? pour vérifier si des œuvres ont été utilisées dans un ensemble de données d'entraînement d'IA (LAION c/ Robert Kneschke, Tribunal régional de Hambourg, 27 septembre 2024)
- Des outils susceptibles d'être utilisés par chacune des parties:
 - les titulaires de droits afin de démontrer l'utilisation des œuvres par l'IA
 - les concepteurs d'outil d'IA souhaitant démontrer l'absence de contrefaçon
- Des outils souvent contraires aux CGU des plateformes et rendus impossibles

EST-CE QUE LE “SCRAPING” EST UN MOYEN LÉGITIME D’ACCÉDER AUX SECRETS D’AFFAIRES ?

- « *Bien que Compulife ait clairement permis au public d'accéder à autant de devis que possible humainement, un robot peut collecter plus de devis qu'un humain ne pourrait pratiquement le faire. Ainsi, bien que l'accès manuel aux devis de la base de données de Compulife ne constitue probablement jamais un moyen illégitime, l'utilisation d'un bot pour collecter une quantité de données autrement infaisable peut bien l'être.* » *Compulife Software Inc. c. Newman*, 959 F.3d 1288, 1314 (11th Cir. 2020)
- Le tribunal de première instance a conclu que les défendeurs avaient détourné le secret d'affaires du plaignant (base de données de devis d'assurance-vie) par des "moyens illégitimes" (une attaque de scraping de données "orchestrée" pour obtenir un avantage concurrentiel).
- La cour d'appel a confirmé, mais : *"Il est important de noter que le scraping et les technologies connexes peuvent être parfaitement légitimes."* *Compulife II*, No. 21-14074 (11th Cir. 1er août 2024)
- L'utilisation d'un bot pour scraper de grandes quantités d'informations disponibles publiquement peut entraîner le détournement de secrets d'affaires



ACCÈS FRAUDULEUX

- *Physicians Interactive c. Lathian Sys., Inc.*, No. CA 03-1193-A, 2003 WL 23018270 (E.D.Va. 5 déc. 2003)

Le plaignant gérait un site web public pour les médecins présentant des produits médicaux et des données pharmaceutiques et stockait également sur le site une base de données confidentielle non accessible au public. Le défendeur a piraté le site web et a extrait les données confidentielles du site.

- Le tribunal a conclu que:
 - « *Il ne fait aucun doute que l'utilisation d'un robot logiciel pour pirater un système informatique et prendre ou copier des informations propriétaires est un moyen illégitime d'obtenir un secret commercial, et constitue donc un détournement en vertu du DTSA* »
 - *le "manque de restriction d'utilisation" du plaignant sur son site web n'a pas automatiquement rendu le piratage légitime.*

VERS LA RECHERCHE D'UN COMPROMIS SUR L'ARTICLE 53(1)(C) RIA

- Mise en place d'un canal de vérification spécifique
- Qui peut demander des informations: seules les « *parties justifiant d'un intérêt légitime* », ce qui implique de fournir :
 - l'identité du titulaire des droits
 - une preuve de propriété et/ou de chaîne de titres
 - une description du contenu protégé
- Quel outil de vérification :
 - aucune demande automatique ou massive pour éviter le *scrapping* et l'ingénierie inverse du contenu exact de l'ensemble de données
 - un engagement à ne pas utiliser la réponse à la demande d'information à d'autres fins que l'exercice des droits d'auteur
 - gestion par le fournisseur d'IA lui-même ou par le Bureau de l'IA

AUTRES PROPOSITIONS

- Proposition faite à la Commission des affaires culturelles de l'Assemblée nationale, dans le cadre des débats concernant les droits d'auteurs et l'intelligence artificielle (19 mars 2025) :
 - renversement de la charge de la preuve en cas de production de l'IA présentant des ressemblances avec une œuvre protégée – les fournisseurs d'IA devront alors fournir une liste exhaustive des données utilisées

CONCLUSION: PROCHAINES ÉTAPES

**Septembre 2024
à avril 2025**

- Atelier de l'Office de l'IA pour rédiger un Code de bonnes pratiques GPAI

2 mai 2025

- Date limite actuelle pour l'adoption du Code de bonnes pratiques GPAI, des lignes directrices sur les modèles GPAI et du modèle pour le résumé des données de formation.

Septembre 2025

- L'étude de l'AIPPI sur la question de droit d'auteur et de l'IA

QUESTIONS?



JONES
DAY®

One Firm Worldwide®

APRAM

ECTA

Caring for a balance of IP rights

Co-organised with:



Sponsored by:



Moderator:
NICOLAS MOREAU
Bignon Lebray, FR

Speakers:
DAVID EL SAYEGH
SACEM, FR

ANGELIQUE DE SIMENCOURT
CHRISTIAN DIOR COUTURE, FR

BENJAMIN BELLAMY
LINAGORA, FR

AI & IP:
A LOVE-HATE RELATIONSHIP?
Afternoon Panel:
Implications of AI and industry perspectives

APRAM-ECTA[®] ETREAT JOINT EVENT (Paris, 3 April 2025, 9:00-17:00)

Q & A

QUESTIONS DE LA SALLE

THANK YOU TO



&



FOR THE CLOSING
COCKTAIL